

# Neutral $B$ -meson mixing from full lattice QCD at the physical point

R. J. Dowdall,<sup>1</sup> C. T. H. Davies<sup>1,2,\*</sup> R. R. Horgan,<sup>1</sup> G. P. Lepage,<sup>3</sup> C. J. Monahan,<sup>4,5,6</sup>  
J. Shigemitsu,<sup>7</sup> and M. Wingate<sup>1†</sup>

(HPQCD Collaboration)<sup>‡</sup>

<sup>1</sup>*DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

<sup>2</sup>*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

<sup>3</sup>*Laboratory for Elementary-Particle Physics, Cornell University, Ithaca, New York 14853, USA*

<sup>4</sup>*Institute for Nuclear Theory, University of Washington, Seattle, Washington 98195, USA*

<sup>5</sup>*Physics Department, College of William and Mary, Williamsburg, Virginia 23187, USA*

<sup>6</sup>*Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, USA*

<sup>7</sup>*Physics Department, The Ohio State University, Columbus, Ohio 43210, USA*



(Received 5 July 2019; published 25 November 2019)

We calculate the bag parameters for neutral  $B$ -meson mixing in and beyond the Standard Model, in full four-flavor lattice QCD for the first time. We work on gluon field configurations that include the effect of  $u$ ,  $d$ ,  $s$ , and  $c$  sea quarks with the highly improved staggered quark (HISQ) action at three values of the lattice spacing and with three  $u/d$  quark masses going down to the physical value. The valence  $b$  quarks use the improved nonrelativistic QCD action and the valence light quarks, the HISQ action. Our analysis was blinded. Our results for the bag parameters for all five operators are the most accurate to date. For the Standard Model operator between  $B_s$  and  $B_d$  mesons we find  $\hat{B}_{B_s} = 1.232(53)$  and  $\hat{B}_{B_d} = 1.222(61)$ . Combining our results with lattice QCD calculations of the decay constants using HISQ quarks from the Fermilab/MILC Collaboration and with experimental values for  $B_s$  and  $B_d$  oscillation frequencies allows determination of the Cabibbo-Kobayashi-Maskawa (CKM) elements  $V_{ts}$  and  $V_{td}$ . We find  $|V_{ts}| = 0.04189(93)$ ,  $|V_{td}| = 0.00867(23)$ , and  $|V_{ts}|/|V_{td}| = 0.2071(27)$ . Our results agree well (within  $2\sigma$ ) with values determined from CKM unitarity constraints based on tree-level processes (only). Using a ratio to  $\Delta M_{s,d}$  in which CKM elements cancel in the Standard Model, we determine the branching fractions  $\text{Br}(B_s \rightarrow \mu^+ \mu^-) = 3.81(18) \times 10^{-9}$  and  $\text{Br}(B_d \rightarrow \mu^+ \mu^-) = 1.031(54) \times 10^{-10}$ . We also give results for matrix elements of the operators  $R_0$ ,  $R_1$ , and  $\tilde{R}_1$  that contribute to neutral  $B$ -meson width differences.

DOI: [10.1103/PhysRevD.100.094508](https://doi.org/10.1103/PhysRevD.100.094508)

## I. INTRODUCTION

The Standard Model (SM) description of neutral  $B_d$  and  $B_s$  oscillations requires knowledge of hadronic parameters derived from the matrix elements of 4-quark operators between  $B_q$  and  $\bar{B}_q$  states. These 4-quark operators come from the effective electroweak Lagrangian at energy scales appropriate to  $B$  physics, and the matrix elements can only be determined by lattice QCD calculations, which are now able to include the full impact of QCD on such hadronic

quantities [1]. The accuracy with which this can be done is the limiting factor in the constraint that can be obtained from the now very precise experimental results on the neutral meson mass difference (seen as an oscillation frequency). In the Standard Model this constraint leads to a determination of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements that accompany the 4-quark operators of the Standard Model. New physics models with extra heavy particles extend the effective Hamiltonian to include additional 4-quark operators. Constraints on the new physics from experiment then need accurate determination of the matrix elements of the new operators. Again this can come only from lattice QCD calculations.

Here we provide the first “second-generation” lattice QCD calculation of the matrix elements of all five  $\Delta B = 2$  operators of dimension six for the  $B_s$  and  $B_d$ . We improve on earlier calculations by working on gluon field configurations generated by the MILC Collaboration that include

\*Christine.Davies@glasgow.ac.uk

†<http://www.physics.gla.ac.uk/HPQCD>

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

$u$ ,  $d$ ,  $s$ , and  $c$  quarks in the sea with  $u/d$  quark masses going down to their physical values. Although this obviates the need for a chiral extrapolation, we also include heavier  $u/d$  quark masses in our set of results so that we can map out the dependence on the light quark mass. The discretization of QCD that we use is fully improved through  $\mathcal{O}(\alpha_s a^2)$  for both the gluons and the light quarks (including all of those in the sea) for which the highly improved staggered quark (HISQ) action is used. For the  $b$  quarks we use improved nonrelativistic QCD, which includes  $\mathcal{O}(\alpha_s)$  corrections to terms at order  $v^4$  (where  $v$  is the heavy quark velocity). By linking this calculation directly to our earlier one for the  $B$ -meson decay constants [2] (that parametrize the amplitude to create a meson from the vacuum) we are able to give results directly for the “bag parameters” associated with each operator and take advantage of the cancellation of a number of systematic effects. The bag parameters (to be defined in Sec. II A) encode the multiplicative factor by which the operator matrix element differs from that expected in the vacuum saturation approximation, which is related to the decay constant. To the extent to which this approximation works (and we will show here that it does work well) we expect the bag parameters to have very little dependence on light quark sea or valence masses and even on the lattice spacing. This enables improvements in accuracy over earlier work along with a much simpler picture of the extrapolation to the physical point.

The first unquenched lattice QCD calculations of the matrix elements for neutral  $B$  meson mixing focused purely on results for the Standard Model operators [3,4] and ratios for  $B_s$  to  $B_d$  [5]. Calculations have also been done in the infinite heavy quark mass limit [6]. More recently calculations of matrix elements for the full set of SM and beyond the Standard Model (BSM) operators have been done [7,8]. The calculations in [7] use the twisted mass formalism for all quarks on gluon field configurations including  $u$  and  $d$  quarks in the sea. An extrapolation of results [renormalized using the regularisation-invariant momentum-subtraction (RI-MOM) scheme] is made from a heavy quark mass in the charm region up to the  $b$  quark mass using ratios with a known infinite mass limit. The calculations in [8] use the Fermilab formalism for the  $b$  quark and the asqtad formalism for the light quarks on gluon field configurations that include  $u$ ,  $d$ , and  $s$  quarks in the sea with the asqtad formalism. Perturbatively renormalized 4-quark operator matrix elements (only) are calculated and so bag parameters must be derived using decay constant results from elsewhere. A very recent result in [9] uses domain-wall quarks on gluon field configurations including  $u$ ,  $d$ , and  $s$  in the sea and extrapolates in heavy quark mass to the  $b$  quark mass for  $B_s$  to  $B_d$  ratios for SM mixing matrix elements and decay constants.

In Sec. II we discuss the 4-quark operators relevant to  $B$  mixing and how they are implemented on the lattice. Section III describes our lattice calculation and results.

We compare our results to previous work in Sec. IV, determine CKM elements  $V_{ts}$  and  $V_{td}$  using experimental results on  $B$ -meson mass differences, determine branching fractions for the rare decays of  $B_d$  and  $B_s$  to  $\mu^+\mu^-$ , and give matrix elements for derived operators that contribute to width differences. Section V gives our conclusions and discusses the prospects for future improvements. Details about our analysis are contained in four appendices: Appendix A on the lattice QCD correlators that we calculate and how we fit them to obtain matrix elements and bag parameters; Appendix B on the chiral perturbation theory fits we use to combine results at physical and unphysical light quark masses; Appendix C on correlations in the uncertainties for our final results and last, with more general applications beyond this analysis, Appendix D on singular value decomposition (SVD) cuts and fitting correlators.

## II. BACKGROUND

### A. Continuum 4-quark operators

Neutral  $B$ -meson mixing occurs at lowest order in the Standard Model through box diagrams involving the exchange of  $W$  bosons and top quarks; see Fig. 1. These box diagrams can be well approximated by an effective Lagrangian expressed in terms of 4-quark operators. Here we will examine all five of the independent local dimension-6 operators that could contribute to  $\Delta B = 2$  processes [8,10]:

$$\begin{aligned} O_1 &= [\bar{\Psi}_b^i (V - A) \Psi_q^j] [\bar{\Psi}_b^j (V - A) \Psi_q^i], \\ O_2 &= [\bar{\Psi}_b^i (S - P) \Psi_q^i] [\bar{\Psi}_b^j (S - P) \Psi_q^j], \\ O_3 &= [\bar{\Psi}_b^i (S - P) \Psi_q^j] [\bar{\Psi}_b^j (S - P) \Psi_q^i], \\ O_4 &= [\bar{\Psi}_b^i (S - P) \Psi_q^i] [\bar{\Psi}_b^j (S + P) \Psi_q^j], \\ O_5 &= [\bar{\Psi}_b^i (S - P) \Psi_q^j] [\bar{\Psi}_b^j (S + P) \Psi_q^i], \end{aligned} \quad (1)$$

where  $V = \gamma_\mu$ ,  $A = \gamma_\mu \gamma_5$ ,  $S = 1$ , and  $P = \gamma_5$ , and sums over  $\mu$  and color indices  $i$  and  $j$  are implicit. In the Standard Model, the most important of these for  $B - \bar{B}$  mixing is  $O_1$ . This operator mixes with  $O_2$  under renormalization. Operators  $O_4$  and  $O_5$  do not appear in the Standard Model, but do arise in various BSM scenarios.

It is conventional to parametrize matrix elements of these operators in terms of “bag parameters,”

$$\begin{array}{c} \bar{b} \text{---} \text{---} \text{---} b \\ \text{---} \text{---} \text{---} \text{---} \\ q \text{---} \text{---} \text{---} \bar{q} \end{array} \quad \equiv \quad \begin{array}{c} \bar{b} \text{---} \text{---} b \\ \text{---} \text{---} \text{---} \\ q \text{---} \text{---} \bar{q} \end{array} + \mathcal{O}(m_b^2/M_Z^2)$$

FIG. 1. An example from the Standard Model of a mechanism that mixes the neutral  $B_q$  and  $\bar{B}_q$ . The amplitude is well approximated by a contact term for matrix elements between  $B_q$ -meson states.

TABLE I. Normalizations  $\eta_i^q(m_b)$  for bag parameters [Eq. (4)]. These are calculated using  $M_{B_s} = 5.3669(2)$  GeV and  $M_{B_d} = 5.2796(2)$  GeV [11];  $\bar{m}_b(\bar{m}_b) = 4.162(48)$  GeV and  $m_b/m_s = 52.55(55)$  [12]; and  $m_s/m_l = 27.18(10)$  [13].

$B_q$	$\eta_1^q$	$\eta_2^q$	$\eta_3^q$	$\eta_4^q$	$\eta_5^q$
$B_s$	2.667	-2.669(62)	0.534 (12)	3.536 (74)	2.068 (25)
$B_d$	2.667	-2.678(62)	0.536 (12)	3.547 (74)	2.071 (25)

$$B_{B_q}^{(i)}(\mu) \equiv \frac{\langle B_q | O_i^q | \bar{B}_q \rangle_{\overline{\text{MS}}}^{(\mu)}}{\eta_i^q(\mu) f_{B_q}^2 M_{B_q}^2}, \quad (2)$$

where here  $\mu$  is the renormalization scale, and  $M_{B_q}$  and  $f_{B_q}$  are the mass and weak decay constant of the  $B_q$  meson:

$$\langle 0 | \bar{\Psi}_q^i \gamma_0 \gamma_5 \Psi_b^i | B_q(\vec{p} = 0) \rangle = f_{B_q} M_{B_q}. \quad (3)$$

The normalization parameter  $\eta_i^q(\mu)$  is chosen so that the bag parameters equal 1 in the “vacuum saturation approximation,” where gluon (and other QCD) exchanges between the initial and final  $\bar{B}_q$  and  $B_q$  are ignored (see [8,10] for more details):

$$\begin{aligned} \eta_1^q &= \frac{8}{3}, \\ \eta_2^q &= -\frac{5}{3} \left( \frac{M_{B_q}}{m_b(\mu) + m_q(\mu)} \right)^2, \\ \eta_3^q &= \frac{1}{3} \left( \frac{M_{B_q}}{m_b(\mu) + m_q(\mu)} \right)^2, \\ \eta_4^q &= 2 \left[ \left( \frac{M_{B_q}}{m_b(\mu) + m_q(\mu)} \right)^2 + \frac{1}{6} \right], \\ \eta_5^q &= \frac{2}{3} \left[ \left( \frac{M_{B_q}}{m_b(\mu) + m_q(\mu)} \right)^2 + \frac{3}{2} \right]. \end{aligned} \quad (4)$$

We use renormalization scale  $\mu = m_b(m_b)$ ; the corresponding values for the normalization factors are given in Table I.

TABLE II. Perturbative coefficients used in Eq. (5) to convert matrix elements of lattice NRQCD-HISQ 4-quark operators into  $\overline{\text{MS}}$  matrix elements. Results are given for the NRQCD valence  $b$ -quark masses (in lattice units) used with our different ensembles. The continuum scheme used is the  $\overline{\text{MS}}_{\text{NDR}}$  scheme of [16] (BBGLN) with  $\mu = m_b$ . The coefficients come from [14], with  $z_{ij} \equiv \rho_{ij} - \zeta_{ij}$  where  $\rho_{ij}$  and  $\zeta_{ij}$  are listed in Tables III and IV of that paper<sup>a</sup>. The perturbative coefficients  $z_{A_0}$  for the temporal axial current [Eq. (A6)] are also listed; these are from [2], which used results from [17].

$am_b$	$z_{11}$	$z_{12}$	$z_{22}$	$z_{21}$	$z_{33}$	$z_{31}$	$z_{44}$	$z_{45}$	$z_{55}$	$z_{54}$	$z_{A_0}$
3.297	-0.472(2)	-0.299(2)	0.440 (2)	0.041 (2)	0.036 (2)	0.092 (2)	0.646 (2)	-0.252(2)	-0.141(2)	0.111(2)	0.024(2)
3.263	-0.469(2)	-0.296(2)	0.438 (2)	0.041 (2)	0.038 (2)	0.091 (2)	0.640 (2)	-0.251(2)	-0.140(2)	0.108(2)	0.022(2)
3.25	-0.469(2)	-0.294(2)	0.438 (2)	0.041 (2)	0.040 (2)	0.091 (2)	0.639 (2)	-0.252(2)	-0.139(2)	0.106(2)	0.022(2)
2.66	-0.429(2)	-0.235(2)	0.394 (2)	0.044 (2)	0.101 (2)	0.080 (2)	0.514 (2)	-0.254(2)	-0.127(2)	0.037(2)	0.006(2)
2.62	-0.427(2)	-0.229(2)	0.388 (2)	0.044 (2)	0.105 (2)	0.080 (2)	0.501 (2)	-0.254(2)	-0.128(2)	0.032(2)	0.001(2)
1.91	-0.296(2)	-0.108(2)	0.340 (2)	0.045 (2)	0.259 (2)	0.053 (2)	0.299 (2)	-0.243(2)	-0.063(2)	-0.084(2)	-0.007(2)

<sup>a</sup>Note that we have corrected two typographical errors, for  $\rho_{21}$  for  $am_b = 2.66$  and  $\zeta_{22}$  for  $am_b = 2.62$ .

The bag parameters provide both computational advantages and physical insights. The leading-order logarithms in chiral perturbation theory, coming from the matrix element of the 4-quark operator and  $f_{B_q}^2$ , partly cancel in the ratio; see Appendix B. In particular, the coefficient of the chiral logarithm from the tadpole diagrams is reduced by a factor of 4. Therefore bag parameters should be less dependent upon the light-quark mass; we find very little mass dependence. Finite-volume effects will be correspondingly reduced. We also find that most of the dependence on lattice spacing cancels. Finally, as we will show, the bag parameters all turn out to be of order one, suggesting that vacuum saturation is a useful approximation. For these reasons, we focus here on bag parameters; values for the matrix elements are easily obtained from the bag parameters given values for the decay constants [2,13].

## B. Lattice QCD 4-quark operators and matching

Matrix elements of the 4-quark operators are regulator dependent, and so we need to convert matrix elements calculated in our simulation (with the lattice regulator) into the corresponding matrix elements for the more conventional  $\overline{\text{MS}}$  scheme. The differences between the two schemes are ultraviolet and so can be calculated using QCD perturbation theory. To the lowest and first order in  $\alpha_s$  the relationship has the form (for  $\mu = m_b$ )

$$\begin{aligned} \langle O_i \rangle_{\overline{\text{MS}}}^{(m_b)} &= (1 + \alpha_s z_{ii}) \langle O_i \rangle_{\text{latt}} + \sum_{j \neq i} \alpha_s z_{ij} \langle O_j \rangle_{\text{latt}} \\ &+ \mathcal{O}\left(\alpha_s^2, \frac{\alpha_s \Lambda_{\text{QCD}}}{m_b}, \alpha_s (a \Lambda_{\text{QCD}})^2\right). \end{aligned} \quad (5)$$

The coefficients  $z_{ij}$  relevant to our simulation were calculated in [14] and are summarized in Table II. The scale for  $\alpha_s$  depends on the lattice spacing; we use the same values for  $\alpha_s$  used in [15] to calculate renormalizations for the axial-vector current that couples to  $B_q$  mesons (see Table IV for the values).

Our lattice analysis uses nonrelativistic QCD (NRQCD) for the  $b$  dynamics. Quarks and antiquarks decouple in NRQCD, and so correspond to separate fields. As a result the lattice version of a 4-quark operator has the form [14]

$$\begin{aligned} & [\bar{\Psi}_b \Gamma_1 \Psi_q] [\bar{\Psi}_b \Gamma_2 \Psi_q] \\ & \rightarrow [\bar{\Psi}_b^{\text{NR}} \Gamma_1 \Psi_q] [\bar{\Psi}_b^{\text{NR}} \Gamma_2 \Psi_q] \\ & + \frac{1}{2m_b} [\nabla \bar{\Psi}_b^{\text{NR}} \cdot \boldsymbol{\gamma} \Gamma_1 \Psi_q] [\bar{\Psi}_b^{\text{NR}} \Gamma_2 \Psi_q] \\ & + \frac{1}{2m_b} [\bar{\Psi}_b^{\text{NR}} \Gamma_1 \Psi_q] [\nabla \bar{\Psi}_b^{\text{NR}} \cdot \boldsymbol{\gamma} \Gamma_2 \Psi_q] + (\Gamma_1 \leftrightarrow \Gamma_2), \quad (6) \end{aligned}$$

where  $\bar{\Psi}_b^{\text{NR}}$  creates a  $b$  quark and  $\bar{\Psi}_b^{\text{NR}}$  annihilates a  $b$  antiquark. These are the lattice operators we use on the right-hand side of Eq. (5). The  $1/m_b$  terms are the  $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$  corrections to the operator in NRQCD.

A complication for operators  $O_2$  and  $O_3$  is the treatment of “evanescent operators.” Our matching results use the  $\overline{\text{MS}}_{\text{NDR}}$  scheme of [16] (BBGLN). These matrix elements are readily converted to the alternative scheme of [18] (BJU) using the following equations [through  $\mathcal{O}(\alpha_s)$ , with  $\mu = m_b$ ] [19,20]:

$$\begin{aligned} O_2|_{\text{BJU}} &= O_2 + \alpha_s(-0.318O_2 - 0.013O_1)|_{\text{BBGLN}}, \\ O_3|_{\text{BJU}} &= O_3 + \alpha_s(0.106O_3 - 0.013O_1)|_{\text{BBGLN}}. \quad (7) \end{aligned}$$

The matching coefficients  $z_{ij}$  from Eq. (5) are plotted against  $am_b$  in Fig. 2. These coefficients are not large and have a relatively benign dependence on the  $b$ -quark mass across the range that we use, although different coefficients behave differently. The diagonal coefficients  $z_{ii}$  are much larger than the corresponding coefficients for the NRQCD-HISQ axial current ( $z_{A_0}$  in Table II), which are unusually small. Note that the only nonzero off-diagonal coefficients are for  $ij$  equal to 12, 21, 31, 45, and 54, and that these tend to be smaller than the diagonal parameters.

It is worth remarking here on the similarities and differences between the perturbative matching we apply here and that used by the Fermilab/MILC Collaborations [8] in their determination of  $B$  mixing matrix elements. They also make use of a perturbative calculation of the matching to  $\mathcal{O}(\alpha_s)$ . They do this after a nonperturbative determination of factors that are needed to remove normalization artifacts from the clover and asqtad actions that they use for heavy and light quarks, respectively, and without which they would have large  $\mathcal{O}(\alpha_s)$  coefficients. We do not need to apply this procedure because the NRQCD and HISQ actions are well behaved in this respect [21]. After applying their nonperturbative procedure, the Fermilab/MILC Collaborations give results for their  $\mathcal{O}(\alpha_s)$  coefficients in Table III of [8]. Their coefficients differ from ours because they are using a different discretization of QCD for both heavy and light quarks.

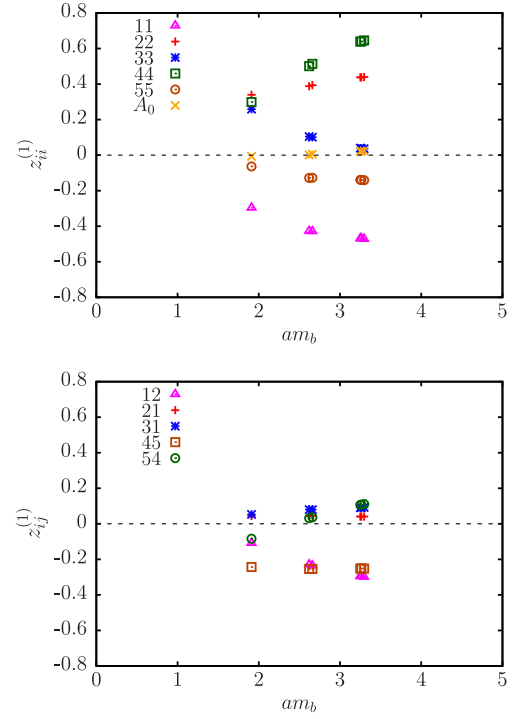


FIG. 2. Coefficients  $z_{ij}$  of  $\mathcal{O}(\alpha_s)$  terms in the matching of lattice NRQCD-HISQ 4-quark operators to the  $\overline{\text{MS}}$  scheme plotted as a function of the bare NRQCD  $b$  quark mass in lattice units. The top plot shows the diagonal coefficients ( $i = j$ ) that enter the renormalization of a given operator; the lower plot shows the off-diagonal coefficients ( $i \neq j$ ) corresponding to the mixing of different operators. See Eq. (5) for the definition of  $z_{ij}$  and Table II for the values. The  $ij$  values are indicated in the key. “ $A_0$ ” refers to the  $\mathcal{O}(\alpha_s)$  coefficient for the renormalization of the temporal axial current ( $z_{A_0}$  in Table II).

However, qualitatively the coefficients show similar behavior in terms of magnitude and dependence on the lattice  $b$  quark mass (given in their case by the parameter  $\kappa'_b$ ).

### III. LATTICE CALCULATION

#### A. Simulations

We use seven ensembles of gluon field configurations recently generated by the MILC Collaboration [22,23]. Details are given in Table III. We use ensembles at three values of the lattice spacing,  $a$ , to control discretization effects and at three values of the light quark mass down to the physical point to map out sea quark mass effects. Discretization effects depend on  $a^2$  and sea quark mass effects are approximately linear, so a range in  $a^2$  of a factor of 3 and in sea light quark mass of a factor of 5 allows us substantial leverage to pin down these effects. The lattice spacing values were determined using the mass splitting between  $\Upsilon$  and  $\Upsilon'$ , as described in [24] where a discussion of systematic errors can be found. The sea quarks use HPQCD’s HISQ action [25] which we have shown to have



TABLE III. Parameters of the gauge ensembles used in this calculation.  $\beta$  is the gauge coupling,  $a_\Upsilon$  is the lattice spacing as determined by the  $\Upsilon(2S - 1S)$  splitting in [24], where the three errors are statistics, NRQCD systematics, and experiment.  $am_l$ ,  $am_s$ , and  $am_c$  are the sea quark masses,  $L \times T$  gives the spatial and temporal extent of the lattices, and  $n_{\text{cfg}}$  is the number of configurations in each ensemble. We use 16 time sources on each configuration to improve statistics.

Set	$\beta$	$a_\Upsilon$ [fm]	$am_l$	$am_s$	$am_c$	$L \times T$	$n_{\text{cfg}}$
1	5.8	0.1474(5)(14)(2)	0.013	0.065	0.838	$16 \times 48$	1000
2	5.8	0.1463(3)(14)(2)	0.0064	0.064	0.828	$24 \times 48$	1000
3	5.8	0.1450(3)(14)(2)	0.00235	0.0647	0.831	$32 \times 48$	1000
4	6.0	0.1219(2)(9)(2)	0.0102	0.0509	0.635	$24 \times 64$	1000
5	6.0	0.1195(3)(9)(2)	0.00507	0.0507	0.628	$32 \times 64$	1000
6	6.0	0.1189(2)(9)(2)	0.00184	0.0507	0.628	$48 \times 64$	1000
7	6.3	0.0884(3)(5)(1)	0.0074	0.037	0.440	$32 \times 96$	1007

small discretization errors even for charm quarks [26–28]. This enables four flavors of quarks to be included in the sea, with masses given in Table III. The  $u$  and  $d$  quark masses are taken to be the same.

The valence  $b$  quarks are implemented using lattice NRQCD [29]. The action is described in detail in [24]. It includes a number of improvements over earlier calculations, in particular one-loop radiative corrections (beyond tadpole improvement [30]) to most of the coefficients of the  $\mathcal{O}(v_b^4)$  relativistic correction terms. The tadpole improvement of the action is done using the Landau gauge link, with  $u_{0L}$  values given in Table IV. This action has been shown to give excellent agreement with experiment in recent calculations of the bottomonium [24,31,32] and  $B$ -meson spectra [33]. The  $b$  quark mass is tuned, giving the values in Table IV, by fixing the spin-averaged kinetic mass of the  $\Upsilon$  and  $\eta_b$  states to experiment [24]. NRQCD breaks down as  $am_b \rightarrow 0$  but all our values of  $am_b$  are substantially larger than 1, where there is no problem.

TABLE IV. Parameters used for the valence quarks.  $am_b$  is the bare  $b$  quark mass in lattice units,  $u_{0L}$  is the Landau link value used for tadpole improvement, and  $am_l^{\text{val}}$  and  $am_s^{\text{val}}$  are the HISQ light and strange quark valence masses. We also tabulate the values from [15] for the running coupling constant  $\alpha_s$  to be used in Eq. (5) for matching lattice 4-quark operators to the continuum. This is in the  $\overline{\text{MS}}$  scheme at scale  $(2/a)$ .

Set	$am_b$	$u_{0L}$	$am_l^{\text{val}}$	$am_s^{\text{val}}$	$\alpha_s$
1	3.297	0.8195	0.013	0.0641	0.346
2	3.263	0.82015	0.0064	0.0636	0.345
3	3.25	0.819467	0.00235	0.0628	0.343
4	2.66	0.834	0.01044	0.0522	0.311
5	2.62	0.8349	0.00507	0.0505	0.308
6	2.62	0.834083	0.00184	0.0507	0.307
7	1.91	0.8525	0.0074	0.0364	0.267

The HISQ valence light quark masses are taken to be equal to the sea mass except on set 4 where there is a slight discrepancy. The  $s$  quark is tuned using the mass of the  $\eta_s$  meson [34], a fictitious pseudoscalar  $s\bar{s}$  state which is not allowed to decay on the lattice. Its properties can be very accurately determined in lattice QCD, and we find  $M_{\eta_s} = 0.6885(22)$  GeV [35]. Values for valence  $s$  masses are given in Table IV and corresponding values of  $M_{\eta_s}$  in lattice units in [35]. We allow for uncertainties from mistuned valence masses in our determination of the physical results.

## B. Simulation results and error budget

We describe the 2-point and 3-point correlators used in our analysis in Appendix A. We use the 2-point correlators to extract the decay constants  $f_{B_q}$ , including the  $1/m_b$  corrections [Eq. (A6)]. We also combine them with the 3-point correlators to calculate lattice matrix elements of the  $\mathcal{O}_n$ . Also in that appendix, we discuss the Bayesian fits used to extract physics from these correlators. Our final results for  $\langle B_q | \mathcal{O}_n | \bar{B}_q \rangle_{\text{latt}} / (f_{B_q} M_{B_q})^2$  are summarized in Table VIII of Appendix A. As discussed in the appendix, this was a blind analysis.

We convert the lattice expectation values into  $\overline{\text{MS}}$  matrix elements using Eq. (5) [divided by  $(f_{B_q} M_{B_q})^2$ ]. Our results are listed in Table V. In addition to the statistical errors from the simulation and the (negligible) errors in the  $z_{ij}$ , we include uncertainties (for each entry in the table) coming from three additional sources:

- (i)  $\mathcal{O}(\alpha_s^2)$ : We estimate this uncertainty to be twice (to be conservative)  $\alpha_s$  times the magnitude of the  $\mathcal{O}(\alpha_s)$  correction we include for each of our three lattice spacings. (These corrections are correlated between configuration sets with similar lattice spacings.)
- (ii)  $\mathcal{O}(\alpha_s \Lambda_{\text{QCD}}/m_b)$ :  $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$  corrections have been measured for the temporal axial-vector current and found to be 5% of the leading-order contribution [2]. This suggests that  $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$  corrections, which are included in our simulations, are 10% for the 4-quark operators. We account for the  $\mathcal{O}(\alpha_s)$  radiative corrections to these terms by adding the following uncertainty to our results:

$$\alpha_s (c_{\alpha_s}^{n,0} + c_{\alpha_s}^{n,1} \delta_a + c_{\alpha_s}^{n,2} \delta_a^2) \frac{\langle \mathcal{O}_n \rangle_{\overline{\text{MS}}}}{(fM)^2}, \quad (8)$$

where each  $c_{\alpha_s}^{n,i} = 0 \pm 0.1$  and

$$\delta_a \equiv (am_b - 2.6)/1.4 \quad (9)$$

allows for variation in the coefficients between the lattice spacings. ( $\delta_a$  is defined to vary from  $-1/2$  to  $1/2$  over our mass range; see [24] for more details.)

TABLE V.  $\overline{\text{MS}}$  matrix elements obtained from Eq. (5) together with simulation data from Table VIII. Results are given for each configuration set and both  $B_s$  (top) and  $B_d$  (bottom) mesons. Values are also given (“phys.”) for our final results at physical quark masses.

Set	$\langle B_s   O_n   \bar{B}_s \rangle_{\overline{\text{MS}}}^{(m_b)} / (f_{B_s} M_{B_s})^2$				
	$O_1$	$O_2$	$O_3$	$O_4$	$O_5$
1	2.10 (21)	-2.14(22)	0.442 (59)	3.56 (41)	1.90 (14)
2	2.16 (21)	-2.20(22)	0.441 (59)	3.73 (43)	2.00 (14)
3	2.14 (21)	-2.18(22)	0.441 (58)	3.69 (41)	1.97 (14)
4	2.20 (15)	-2.19(16)	0.443 (48)	3.76 (28)	2.017 (97)
5	2.15 (14)	-2.16(16)	0.432 (47)	3.64 (27)	1.948 (91)
6	2.20 (14)	-2.20(16)	0.445 (48)	3.73 (27)	1.990 (93)
7	2.19 (10)	-2.19(12)	0.443 (37)	3.70 (16)	1.976 (93)
Phys.	2.168 (93)	-2.18(10)	0.436 (29)	3.65 (15)	1.945 (76)

Set	$\langle B_d   O_n   \bar{B}_d \rangle_{\overline{\text{MS}}}^{(m_b)} / (f_{B_d} M_{B_d})^2$				
	$O_1$	$O_2$	$O_3$	$O_4$	$O_5$
1	2.07 (21)	-2.12(22)	0.438 (59)	3.57 (42)	1.89 (14)
2	2.10 (22)	-2.13(22)	0.421 (60)	3.77 (44)	2.04 (15)
3	2.06 (21)	-2.10(21)	0.398 (58)	3.77 (43)	1.98 (14)
4	2.20 (16)	-2.15(16)	0.403 (48)	3.93 (30)	2.11 (11)
5	2.16 (15)	-2.06(15)	0.396 (48)	3.70 (28)	1.965 (98)
6	2.11 (16)	-2.17(17)	0.447 (52)	3.87 (30)	2.04 (11)
7	2.20 (12)	-2.14(13)	0.413 (39)	3.83 (19)	2.03 (11)
Phys.	2.15 (11)	-2.06(11)	0.400 (30)	3.82 (18)	2.015 (92)

- (iii)  $\mathcal{O}(\alpha_s(a\Lambda_{\text{QCD}})^2, (a\Lambda_{\text{QCD}})^4)$ : The NRQCD and HISQ actions we use in the simulation are highly corrected. In particular, there are no tree-level  $a^2$  errors in either. We account for  $a^2\alpha_s$  errors by adding the following uncertainty to our results:

$$\begin{aligned} & \alpha_s(a\Lambda_{\text{QCD}})^2 (c_{a^2}^{n,0} + c_{a^2}^{n,1} \delta_a + c_{a^2}^{n,2} \delta_a^2) \frac{\langle O_n \rangle_{\overline{\text{MS}}}}{(fM)^2} \\ & + (a\Lambda_{\text{QCD}})^4 (c_{a^4}^{n,0} + c_{a^4}^{n,1} \delta_a + c_{a^4}^{n,2} \delta_a^2) \frac{\langle O_n \rangle_{\overline{\text{MS}}}}{(fM)^2}, \end{aligned} \quad (10)$$

where each  $c_{a^2}^{n,i} = 0 \pm 1$ , each  $c_{a^4}^{n,i} = 0 \pm 1$ ,  $\Lambda_{\text{QCD}} = 0.5 \text{ GeV}$  is the QCD scale, and again the  $\delta_a$  terms allow for variation between different lattice spacings.

The final entries (“phys.”) in both parts of Table V are our final results at the physical values of the light-quark masses after our chiral fit. We use chiral perturbation theory to combine the values obtained on the different configuration sets with different light quark masses; see Appendix B for details. Figure 3 compares our final values for  $B_s$  mesons

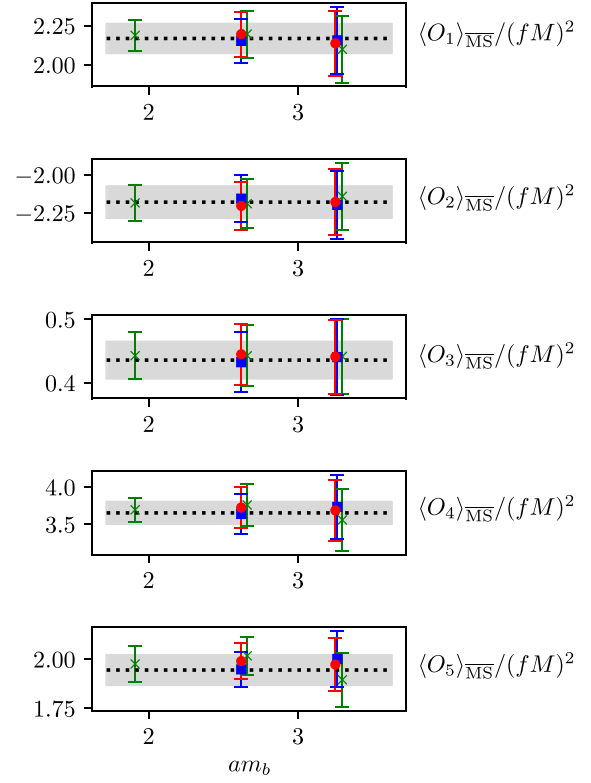


FIG. 3. Comparison of the  $\langle B_s | O_n | \bar{B}_s \rangle_{\overline{\text{MS}}}^{(m_b)} / (f_{B_s} M_{B_s})^2$  values from individual configurations sets (colored data points) with the final extrapolated values (gray bands and dotted lines) for each 4-quark operator. Errors shown include correlated uncertainties from operator normalization and lattice spacing effects as discussed in the text. The data are plotted versus  $am_b$ , falling into three groups corresponding to lattice spacings of 0.09, 0.12, and 0.15 fm. Results are shown for three different values of light-quark mass  $m_l \equiv (m_u + m_d)/2$  corresponding to  $m_l/m_s = 1/5$  (green,  $\times$  s),  $m_l/m_s = 1/10$  (blue, boxes), and the physical mass (red, circles). The dotted lines show the extrapolated values, while the gray bands show the  $\pm 1\sigma$  uncertainty in those values. The analogous figure for  $B_d$  mesons is very similar.

with the results from individual configuration sets. These plots show that the dependence on lattice spacing and light-quark mass is negligible compared with our uncertainties. The analogous plot for  $B_d$  mesons is very similar.

Adding uncertainties to the lattice results to allow for operator normalization and lattice spacing effects, as we have done above, is equivalent to including them in our fit function with the coefficients treated as fit parameters; see the appendix of [36]. The uncertainties that are included are correlated between lattice results on different sets through these coefficients. Figure 3 shows that, for example, the lattice spacing effects that we allow for through Eq. (10) are overestimates of what is seen in the results, since the variation with lattice spacing of the central values is much smaller than the uncertainties shown on the coarser lattices. A further test of this is that omitting the results from the smallest lattice spacing

TABLE VI.  $\overline{\text{MS}}$  bag parameters [Eq. (2) with  $\mu = m_b$ ] for the five 4-quark operators. Results are given for both  $B_s$  and  $B_d$  mesons, and for the ratios of bag parameters.

	$B_{B_q}^{(1)}(m_b)$	$B_{B_q}^{(2)}(m_b)$	$B_{B_q}^{(3)}(m_b)$	$B_{B_q}^{(4)}(m_b)$	$B_{B_q}^{(5)}(m_b)$
$B_s$	0.813 (35)	0.817 (43)	0.816 (57)	1.033 (47)	0.941 (38)
$B_d$	0.806 (40)	0.769 (44)	0.747 (59)	1.077 (55)	0.973 (46)
$B_s/B_d$	1.008 (25)	1.063 (24)	1.092 (34)	0.959 (21)	0.967 (23)

(set 7) shifts our final central values by less than half a standard deviation and often much less.

Finally we convert our final results into bag parameters using Eq. (2). The bag parameters are listed in Table VI. Despite the wide variation in values for  $\langle O_n \rangle / (fM)^2$ , the bag parameters are within 30% of 1. This shows that the vacuum saturation approximation can be of some utility.

Figure 4 compares our final results for ratios of bag parameters  $B_{B_s}^{(n)}/B_{B_d}^{(n)}$  with results from the different

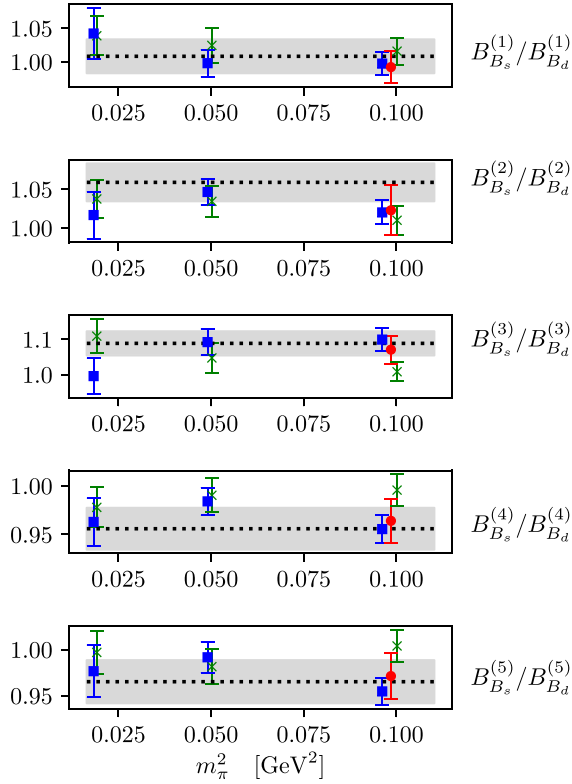


FIG. 4. Comparison of the ratio of bag parameters  $B_{B_s}^{(n)}/B_{B_d}^{(n)}$  from individual configuration sets (colored data points) with the final extrapolated values (gray bands and dotted lines) for each 4-quark operator. The data are plotted versus  $m_\pi^2$ , falling into three groups corresponding approximately to the physical value, 2.7 times the physical value, and 5.4 times the physical value. Results are shown for three different lattice spacings corresponding approximately to 0.15 fm (green,  $\times$  s), 0.12 fm (blue, boxes), and 0.09 fm (red, circles). The dotted lines show the extrapolated values, while the gray bands show the  $\pm 1\sigma$  uncertainty in those values.

TABLE VII. Percent errors coming from different sources for the  $B_s$  meson's bag parameters  $B_{B_s}^{(n)}$  and  $B_{B_s}^{(1)}/B_{B_d}^{(1)}$  (Table VI). The total error for each quantity is also shown. The error budgets for the  $B_d$  meson's bag parameters are very similar. Systematic errors from finite-volume, QED, and strong-isospin breaking effects are estimated to be below 0.1% and hence negligible in Appendix B 5.

	$B_{B_s}^{(1)}$	$B_{B_s}^{(2)}$	$B_{B_s}^{(3)}$	$B_{B_s}^{(4)}$	$B_{B_s}^{(5)}$	$B_{B_s}^{(1)}/B_{B_d}^{(1)}$
Lattice data	1.4	1.4	1.5	1.6	1.5	1.5
$\eta_i^q$	0.0	2.3	2.3	2.1	1.2	0.0
$\alpha_s^2$ terms	2.1	2.9	5.2	1.9	1.5	0.1
$\alpha_s \Lambda_{\text{QCD}}/m_b$ terms	2.9	2.8	2.9	2.8	2.7	0.0
$(a\Lambda_{\text{QCD}})^{2n}$ terms	1.8	1.9	2.3	1.5	1.8	0.1
$m_l$ extrapolation	0.4	0.4	0.7	0.5	0.4	1.9
Total	4.3	5.3	7.0	4.6	4.1	2.5

configuration sets. Results are plotted versus the value of  $m_\pi^2$  used in each simulations. Again there is very little variation with quark mass, with all ratios within 5% of 1. Our final results are shifted by less than half a standard deviation if we omit the data with the largest pion masses, and have errors that are 10%–15% larger.

The error budgets for the  $B_s$  bag parameters are shown in Table VII. The dominant source of error comes from uncalculated terms in perturbation theory ( $\alpha_s^2$  and  $\alpha_s \Lambda_{\text{QCD}}/m_b$  terms). The sensitivity to these terms depends on the operator. For example, it is particularly high for  $O_3$ , because matrix elements for  $O_3$  are a lot smaller than those of  $O_1$  [see Eq. (4)] which are mixed in by Eq. (5). The error budgets for  $B_d$  mesons are almost identical to those for  $B_s$ , but have twice as large a contribution from statistical uncertainties in the lattice data. Almost all of the uncertainties, and some of the statistical errors, cancel in ratios of  $B_s$  to  $B_d$  meson bag parameters.

Matrix elements of the 4-quark mixing operators can be obtained from the ratios in Table V given values for the decay constants and masses. Note that the corresponding bag parameters for  $O_{2\dots 5}$  have larger fractional errors than the ratios, and so should not be used for this purpose. The larger errors result from uncertainties due to the factors  $\eta_i^q$  in the bag-parameter definition [see Table VII and Eq. (2)].

## IV. DISCUSSION

### A. Comparison to previous results

Our results for the bag parameters for all five SM and BSM operators given in Table VI are more accurate than previous lattice QCD results. This is for a number of reasons:

- (i) We work directly with the bag parameters rather than the 4-quark operator matrix elements. The bag parameters are expected from chiral perturbation theory to have little dependence on valence and sea quark masses (see Appendix B). This expectation is

borne out in our results and means that we are easily able to combine results at both unphysical and physical light quark masses.

- (ii) We have results for the physical light quark mass at two values of the lattice spacing improving control of the chiral extrapolation.
- (iii) The gluon field configurations that we use include the effect of  $u$ ,  $d$ ,  $s$ , and  $c$  quarks in the sea and so we do not have an uncertainty associated with missing flavors of sea quarks (the Fermilab/MILC Collaboration includes a 2% uncertainty in their 4-quark operator matrix elements from missing  $c$  in the sea [8]).

Figure 5 shows a comparison of our bag parameters for the  $B_s$  meson to those from [7,8] (and also, for  $O_1$ , to [4]).

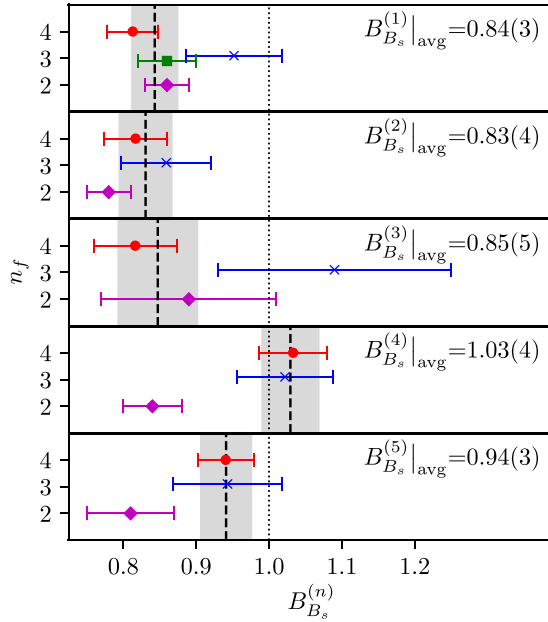


FIG. 5. A comparison of our results (red filled circles at  $n_f = 4$ ) to previous lattice QCD values for the  $B_s$  bag parameters  $B_{B_s}(m_b)$  in the  $\overline{\text{MS}}$  scheme for all five SM and BSM operators. Previous results come from the Fermilab/MILC Collaboration on  $n_f = 3$  gluon field configurations (blue crosses) [8] and the ETM Collaboration on  $n_f = 2$  gluon field configurations (purple filled diamonds) [7]. Note that the ETM results for  $O_4$  and  $O_5$  have been converted to the definition of the bag parameter given in Eq. (4). The filled green square at  $n_f = 3$  for the  $O_1$  operator comes from an earlier HPQCD calculation using NRQCD  $b$  quarks [4]. The  $n_f = 2$  results are missing  $s$  sea quarks, whose impact cannot be estimated perturbatively (and no uncertainty is included for this in the error bars). It is therefore unclear what level of agreement to expect between these results and those for  $n_f = 3$  and 4. Since we do not expect missing  $c$  in the sea to have a significant impact on the bag parameters [8] we can meaningfully compare  $n_f = 3$  and  $n_f = 4$ . The grey bands are the weighted average of our new results with those of [8], and the average value of the bag parameter  $B_{B_s}^{(n)}(m_b)$  for each operator  $O_n$  is indicated in that panel. We include a vertical line at value 1.0 for comparison to the vacuum saturation approximation.

The results from [7] include only  $u$  and  $d$  quarks in the sea and the uncertainty does not include an estimate of the impact of missing  $s$  sea quarks. It is therefore not clear whether we should expect agreement between these  $n_f = 2$  results and our  $n_f = 4$  results. The fact that the  $n_f = 2$  purple diamonds from the ETM Collaboration are around 20% below our results for  $O_4$  and  $O_5$  is reminiscent of what is seen in kaon mixing. ETM uses the RI-MOM renormalization scheme for the purple diamonds, and it has been shown in kaon mixing [37] that the use of the RI-MOM scheme (rather than RI-SMOM) for the equivalent 4-quark operators has large systematic errors that push down the value of the bag parameter. This may then be the main reason (rather than a difference of  $n_f$ ) for the discrepancy with our results for  $O_4$  and  $O_5$ , but more work would be needed to be sure of this.

The  $n_f = 3$  and  $n_f = 4$  results should be comparable because the impact of missing  $c$  quarks in the sea on the bag parameters is expected to be very small [8]. Our new results agree within  $2\sigma$  in each case with [8] but in every case are more accurate. The largest discrepancy is for  $B_{B_s}^{(1)}$  at  $1.9\sigma$ .

The weighted average of our  $n_f = 4$  results and the  $n_f = 3$  results from [8] is given by the grey band in the figure and the value of that average is given in each panel. We assume no correlations, here and subsequently, between our results and those of [8] because they use different actions for both the  $b$  quark and the light quarks and different gluon field configurations (with a different sea quark action and generated with a different Monte Carlo updating algorithm).

Figure 6 shows a comparison of the ratio of bag parameters for  $B_s$  to  $B_d$  for each operator for our new results and those of [8]. Our new results are a lot more accurate, with 2%–3% total uncertainty. All of the ratios are very close to 1, but there is a sign of a systematic trend for the ratio for  $O_2$  and  $O_3$  to be above 1 and for  $O_4$  and  $O_5$  below 1. This is not visible in the results of [8] but does start to emerge with the improved accuracy of our results. This is in general agreement with the results from using sum rules in [38,39]. We also include in Fig. 6 results for  $O_1$  from HPQCD [4] and RBC/UKQCD [9]. The RBC/UKQCD result has a 1% uncertainty.

## B. Derived quantities

Our results for the bag parameters can be combined with results for the  $B$  and  $B_s$  decay constants to give values for the 4-quark operator matrix elements using Eq. (B1) and our results in Table V. For this we use the most accurate current lattice QCD results obtained on gluon field configurations including  $u/d$ ,  $s$ , and  $c$  quarks in the sea. These have been obtained by the Fermilab/MILC Collaboration using the HISQ action for all quarks [13]. This “heavy-HISQ” approach, pioneered by HPQCD [40,41], uses pseudoscalar meson 2-point correlators that combine heavy



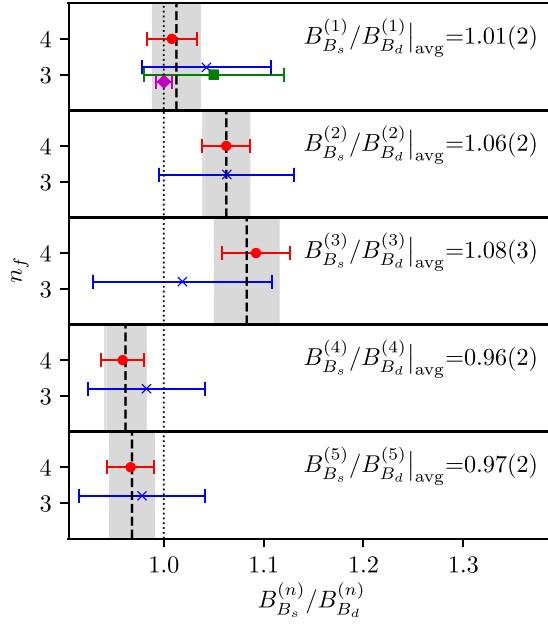


FIG. 6. A comparison of our results (red filled circles at  $n_f = 4$ ) to previous lattice QCD values for the ratio of  $B_s$  to  $B_d$  bag parameters for all five SM and BSM operators. Previous results come from the Fermilab/MILC Collaboration on  $n_f = 3$  gluon field configurations (blue crosses) [8] using their quoted correlations to reconstruct the ratio. Since we do not expect missing  $c$  in the sea to have a significant impact on the bag parameters [8] we can meaningfully compare  $n_f = 3$  and  $n_f = 4$ . The grey bands are the weighted average of these two sets of results and the average value for each operator is indicated in that panel. For  $O_1$  at  $n_f = 3$  we also show previous results from HPQCD (green filled square) using NRQCD  $b$  quarks [4] and RBC/UKQCD (purple filled diamond) using domain-wall quarks with masses of  $m_c$  and above and extrapolating results to the  $b$  quark mass [9]. We include a vertical line at value 1.0 to make clear which ratios are above, and which below, this value.

and light quark propagators calculated with multiple heavy quark masses,  $am_h$ , at multiple values of the lattice spacing.  $m_h$  reaches the  $b$  quark mass for  $am_h < 1$  for lattice spacing values  $a < 0.045$  fm. Since the HISQ action has very small discretization errors by design, a fit to the  $m_h$  and  $a$  dependence is possible that allows the continuum  $m_h$  dependence of the decay constant to be reconstructed. It can then be evaluated at the  $b$  quark mass to enable the  $B$  and  $B_s$  decay constants to be determined. Note that the correlators can be absolutely normalized in this case and so there is no normalization uncertainty.

Fermilab/MILC obtain the values  $f_{B_d} = 0.1905(13)$  GeV,  $f_{B_s} = 0.2307(13)$  GeV, and  $f_{B_s}/f_{B_d} = 1.2109(41)$ .<sup>1</sup> Note that we use the decay constant for the neutral  $B_d$  meson (not the  $B_u$ ), which is the appropriate choice here. Our bag

parameters are calculated for a light quark  $l$  corresponding to the average of  $u$  and  $d$ . Our results show (comparing those for  $B_s$  with those for  $B_d$ ) that any difference between bag parameters for  $B_l$  and  $B_d$  will be much smaller than our uncertainties. This is not true for the decay constants, where the differences are significant [13].

For the SM phenomenology to be determined from our results for the matrix elements of  $O_1$  it is convenient to convert our results from the  $\overline{\text{MS}}_{\text{NDR}}$  scheme to the renormalization-group-invariant quantities  $\hat{B}_{B_q}^{(1)}$ . The conversion is given by

$$\hat{B}_{B_q}^{(1)} = c_{\text{RGI}} B_{B_q}^{(1)}(m_b). \quad (11)$$

The matching factor  $c_{\text{RGI}}$  is calculated to two loops in perturbative QCD, and we take  $c_{\text{RGI}} = 1.5158(36)$  [8]. This corresponds to the result for  $n_f = 5$  active flavors in the sea and  $\bar{\alpha}_s(M_Z) = 0.1185(6)$ . Our bag parameters are obtained at scale  $m_b$  for four flavors of quarks in the sea. The impact of missing  $b$  quarks in the sea, however, should be negligible both for the bag parameters and the resulting 4-quark operator matrix elements. A power-counting estimate of such effects would give a relative contribution of  $\alpha_s(\Lambda_{\text{QCD}}/2m_b)^2$ , which is below 0.1%.

Our results for the RGI bag parameters for  $O_1$  are then

$$\begin{aligned} \hat{B}_{B_s}^{(1)} &= 1.232(53), \\ \hat{B}_{B_d}^{(1)} &= 1.222(61), \\ \frac{\hat{B}_{B_s}^{(1)}}{\hat{B}_{B_d}^{(1)}} &= 1.008(25). \end{aligned} \quad (12)$$

The ratio of RGI bag parameters is of course the same as that of the  $\overline{\text{MS}}$  bag parameters. Combined with the decay constant results from [13] we obtain

$$\begin{aligned} f_{B_s} \sqrt{\hat{B}_{B_s}^{(1)}} &= 0.2561(57) \text{ GeV}, \\ f_{B_d} \sqrt{\hat{B}_{B_d}^{(1)}} &= 0.2106(55) \text{ GeV}, \\ \xi &= 1.216(16), \end{aligned} \quad (13)$$

where  $\xi$  is the ratio of the two results above it. We form  $\xi$  by combining the result for  $f_{B_s}/f_{B_d}$  from [13] with our results for  $\hat{B}_{B_s}^{(1)}/\hat{B}_{B_d}^{(1)}$ , taking advantage of the correlations that reduce uncertainties in each of these ratios. Note that in combining the decay constant and bag parameter results we add relative uncertainties in quadrature. We expect no significant correlation between the two sets of results because they use a different heavy quark action and, even though both results use  $n_f = 2 + 1 + 1$  gluon field configurations, there is little overlap in the ensembles used.

<sup>1</sup>Our results obtained on  $n_f = 2 + 1 + 1$  gluon field configurations from NRQCD-HISQ calculations [2,42] agree with these numbers but are less accurate.

The error budgets in the two cases show that the key sources of uncertainty are not the same. The uncertainties in the combinations above are dominated by the uncertainties in our bag parameters and their ratio in Eq. (12) because the decay constant results are now so accurate.

Figure 7 compares our new result for  $f_{B_s} \sqrt{\hat{B}_{B_s}}$  from Eq. (13) to previous lattice QCD results on  $n_f = 3$  gluon field configurations from Fermilab/MILC [8] and HPQCD [4]. The Fermilab/MILC results include an uncertainty for missing  $c$  in the sea in their calculation. The difference between the central value of our new result and that of Fermilab/MILC is  $1.8\sigma$ . Because the systematic uncertainties are correlated between our results and those of [4] we do not include the previous HPQCD results in the new lattice QCD  $n_f = 3/n_f = 4$  average, shown by the grey band in Fig. 7. The average value is shown above the grey band.

Figure 8 compares lattice QCD results for the ratio  $\xi$  defined in Eq. (13) on  $n_f = 3$  gluon field configurations with our new result here using  $n_f = 4$ . There is good agreement between the lattice QCD results with the most recent (including our new result here) having total uncertainties at the level of 1.5%. The result of averaging our new result with that of [8] (both results being obtained at the physical  $b$  quark mass) is given by the grey band with the average value quoted above it.

### C. $\Delta M$

The phenomenon of neutral  $B$ -meson oscillations is now well-established experimentally (for recent results see [43–51]), with an oscillation frequency that is set by the mass difference between the two eigenstates.

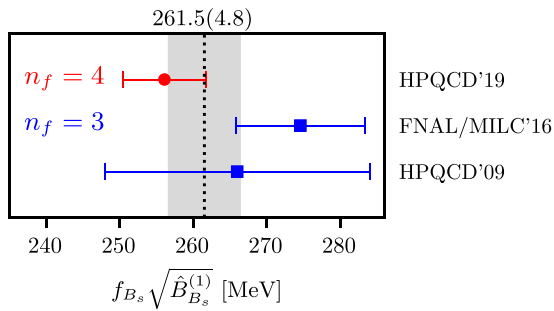


FIG. 7. A comparison of our results (red filled circles at  $n_f = 4$ ) to previous lattice QCD values for the combination of decay constant and square root of bag parameter  $f_{B_s} \sqrt{\hat{B}_{B_s}^{(1)}}$ . Previous results (blue filled squares) come from the Fermilab/MILC Collaboration [8] and from HPQCD [4] on  $n_f = 3$  gluon field configurations. The Fermilab/MILC results include a 1% uncertainty for missing  $c$  in the sea. The grey band is the weighted average of our new results and those of [8] and the new lattice QCD average value is quoted at the top.

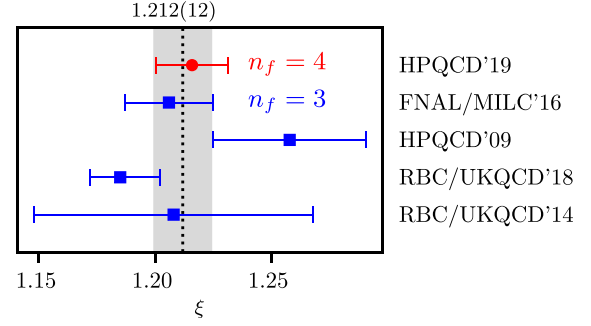


FIG. 8. A comparison of our results (red filled circles at  $n_f = 4$ ) for  $\xi$ , defined in Eq. (13), to previous lattice QCD values for  $n_f = 3$  (filled blue squares). Previous results come from the Fermilab/MILC Collaboration [8] and from HPQCD [4] using calculations at the physical  $b$  quark mass. Results are also shown from RBC/UKQCD using domain-wall quarks and extrapolating to the  $b$  from the  $c$  quark region and above [9] and using static (infinitely massive)  $b$  quarks [6]. The grey band is the weighted average of our new results and those of [8] with the result for the average quoted above it.

The current experimental average values [11] for the  $B_s$  and  $B_d$  systems are

$$\begin{aligned}\Delta M_{s,\text{expt}} &= 17.757(21) \text{ ps}^{-1}, \\ \Delta M_{d,\text{expt}} &= 0.5065(19) \text{ ps}^{-1},\end{aligned}\quad (14)$$

combining statistical and systematic errors in quadrature. In the SM  $\Delta M$  is given by

$$\Delta M_q = \frac{G_F^2 M_W^2 M_{B_q}}{6\pi^2} S_0(x_t) \eta_{2B} |V_{tq}^* V_{tb}|^2 f_{B_q}^2 \hat{B}_{B_q}^{(1)}. \quad (15)$$

Here  $S_0$  is the Inami-Lim function [52] which describes electroweak corrections and has argument  $x_t = m_t^2/M_W^2$ . The top quark mass to be used here is in the  $\overline{\text{MS}}$  scheme,  $\bar{m}_t(\bar{m}_t)$  [53]. Taking the current average [11] of direct experimental measurements [54–56] of the top quark mass [172.9(4) GeV] as the pole mass, gives  $\bar{m}_t(\bar{m}_t) = 163.07(38)$  GeV using the 4-loop expressions in [57]. Evaluating the Inami-Lim function then gives  $S_0(4.116(19)) = 2.313(8)$ . The QCD correction factor,  $\eta_{2B}$ , is given at next-to-leading order in [58]. We take  $\eta_{2B} = 0.55210(62)$  [13], again calculated with  $n_f = 5$ .

The CKM elements  $V_{tq}$  and  $V_{tb}$  can be derived in the SM by assuming that the CKM matrix is unitary and determining other CKM elements in the same rows or columns from the comparison of theory and experiment [59–62]. For Eq. (15) it is important to use values for  $V_{tq}$  that did not include  $\Delta M_q$  itself in their determination. So we use the results from CKMfitter for the case where only tree-level processes were used in the determination. This gives [61]

$$\begin{aligned}
|V_{ts}|_{\text{CKMfitter, tree}} &= (41.69^{+0.39}_{-1.45}) \times 10^{-3}, \\
|V_{td}|_{\text{CKMfitter, tree}} &= (9.08^{+0.23}_{-0.45}) \times 10^{-3}, \\
|V_{td}/V_{ts}|_{\text{CKMfitter, tree}} &= 0.2186^{+0.0049}_{-0.0059}, \\
|V_{tb}|_{\text{CKMfitter, tree}} &= 0.999093^{+0.000064}_{-0.000018}. \quad (16)
\end{aligned}$$

The ratio  $|V_{td}/V_{ts}|_{\text{CKMfitter, tree}}$  is derived from the CKMfitter results for  $A$ ,  $\lambda$ ,  $\bar{\rho}$ , and  $\bar{\eta}$  using the formulas in [59]. The central value differs slightly from the ratio of the two numbers above.

The final terms in Eq. (15) parametrize the hadronic contribution to  $\Delta M$  through the matrix element of the appropriate 4-quark operator,  $\mathcal{O}_1$ . Our results for  $f_{B_q}^2 \hat{B}_{B_q}^{(1)}$  are given in Eq. (13).

Putting all these pieces together we obtain predictions for the mass differences for neutral  $B_s$  and  $B_d$  eigenstates of

$$\begin{aligned}
\Delta M_{s,\text{SM}} &= 17.59^{(+0.33}_{-1.22})(0.78) \text{ ps}^{-1}, \\
\Delta M_{d,\text{SM}} &= 0.555^{(+28}_{-55})(29) \text{ ps}^{-1}, \\
\left(\frac{\Delta M_d}{\Delta M_s}\right)_{\text{SM}} &= 0.0318^{(+14}_{-17})(8), \quad (17)
\end{aligned}$$

where the first error in each case is from the CKM matrix elements and the second error is primarily from the lattice analyses. These results agree well with the experimental values from Eq. (14)—the largest discrepancy is  $1.7\sigma$  for the ratio of  $\Delta M$  values—but they have much larger uncertainty.

#### D. $V_{ts}$ and $V_{td}$

Because the experimental values for  $\Delta M_q$  are so accurate, a better approach to understanding the implications of our improved lattice QCD results for the relevant hadronic matrix elements is to turn the analysis of the previous subsection on its head. That is, to use our results and the experimental values for  $\Delta M_q$  to determine values for  $|V_{ts}|$  and  $|V_{td}|$  from Eq. (15) [taking a value for  $V_{tb}$  from Eq. (16) [61]].  $|V_{ts}|$  and  $|V_{td}|$  obtained this way can then be compared to other determinations that make use of CKM unitarity as a test of that unitarity.

The ratio of  $|V_{ts}|$  to  $|V_{td}|$  can be obtained more accurately than the separate CKM elements because this makes use of the hadronic parameter  $\xi$  [Eq. (13)] in which a lot of the lattice QCD uncertainties cancel (see Sec. IV A).

Our results are

$$\begin{aligned}
|V_{td}| &= 0.00867(23), \\
|V_{ts}| &= 0.04189(93), \\
|V_{td}|/|V_{ts}| &= 0.2071(27). \quad (18)
\end{aligned}$$

Figure 9 plots the  $\pm 1\sigma$  constraints on  $|V_{td}|$ ,  $|V_{ts}|$  and their ratio from our results as the dark grey lozenge. Results

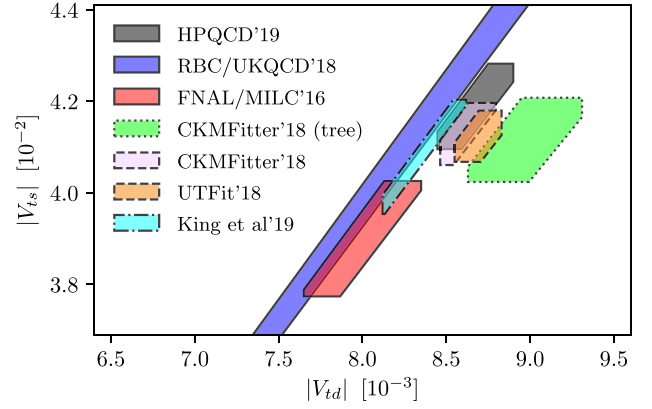


FIG. 9. A comparison of  $\pm 1\sigma$  constraints on  $V_{ts}$  and  $V_{td}$  from experimental results on  $B_s$  and  $B_d$  oscillation frequencies compared to SM calculations. This is an update of Fig. 7 in [39] to include the results presented here. The lattice QCD constraints shown come from the following: this paper, dark grey; [8], red; [9], light blue,  $|V_{ts}|/|V_{td}|$  ratio only. The light blue lozenge is from sum rules [39]. The lozenges with dashed boundaries include a full unitarity triangle fit: light pink is from CKMfitter [59,61] and orange from UTFit [60,62]. The green lozenge with dotted boundary is the result of a unitarity triangle fit for tree-level processes only from CKMfitter.

determined by other lattice QCD calculations [8,9] are also shown along with a recent determination using sum rules [39]. Also shown as light pink and orange lozenges are results from fits to the CKM unitarity triangle using results from many different processes [61,62]. Particularly relevant here is the green lozenge which results from a unitarity triangle fit that includes tree-level processes only [61], and therefore not  $B_s/B_d$  oscillations. Tension between results derived from  $\Delta M_q$  (as here) and the results derived from tree-level processes and unitarity would imply the existence of new physics in loop processes.

The Fermilab/MILC results (red lozenge in Fig. 9) highlighted an approximately  $2.0\sigma$  tension between their values for  $V_{ts}$  and  $V_{td}$  and those from unitarity fits. See [63,64] for examples of the possible implications of this.

Our results show no such tension. Our values for  $V_{ts}$  and  $V_{td}$  separately agree with the {CKMfitter, tree} results in Eq. (16) within  $1\sigma$  and the difference in the ratio amounts to  $1.8\sigma$ . This limits the scope for new physics in loop-induced processes. However, our ratio for  $|V_{td}|/|V_{ts}|$  joins the systematic trend of the previous results shown in Fig. 9 in being below that of {CKMfitter, tree}.

#### E. $B_q \rightarrow \mu^+ \mu^-$ decay

The rare decays  $B_q \rightarrow \mu^+ \mu^-$  have very small branching fractions in the SM since they proceed through  $W$  box diagrams and  $Z$  penguins and are helicity suppressed. New physics might then be seen if the experimental and SM branching fractions can be determined to be different to sufficient accuracy.

The hadronic parameter that enters the SM branching fraction is the  $B_q$  meson decay constant [65] but it appears along with the CKM elements  $|V_{tq}^* V_{tb}|$ . The uncertainty in the value of the appropriate CKM element is now the largest uncertainty in the value of the SM branching fraction [13].

An alternative method for determining the branching fraction is to take a ratio to  $\Delta M$  [66]. In the SM (and extensions with minimal flavor violation) the CKM elements cancel out of this ratio. The decay constant also cancels and the hadronic parameter that remains in the ratio is the bag parameter.

The formula for the time-averaged branching fraction [67], as measured in the experiment, is then given in the SM by

$$\frac{\text{Br}(B_q \rightarrow \ell^+ \ell^-)}{\Delta M_q} = \frac{3G_F^2 M_W^2 m_\ell^2}{\pi^3} \tau_{B_q^H} \sqrt{1 - \frac{4m_\ell^2}{M_{B_q}^2}} \frac{|C_A(\mu_b)|^2}{S_0(x_t) \eta_{2B} \hat{B}_{B_q}^{(1)}}. \quad (19)$$

Here  $C_A(\mu_b)$  includes electroweak and QCD corrections and is given for  $\mu_b = 5$  GeV in [65]. We use  $C_A(\mu_b) = 0.4694(36)$  [8]. The lifetime  $\tau_{B_q^H}$  that appears in this formula is that of the heavy neutral eigenstate [67]. For the  $B_d$  this can be taken as the average lifetime, 1.520(4) ps [68,69], but for the  $B_s$  there is a measured difference of lifetimes and the heavy eigenstate has the longer lifetime, 1.615(9) ps [68,69]. Values for  $S_0(x_t)$  and  $\eta_{2B}$  are given in Sec. IV D.

Using our results for the bag parameters for  $\mathcal{O}_1$  given in Eq. (12) and the experimental values for  $\Delta M$  in Eq. (14) we obtain the following values for the branching fractions:

$$\begin{aligned} \text{Br}(B_s \rightarrow \mu^+ \mu^-) &= 3.81(18) \times 10^{-9}, \\ \text{Br}(B_d \rightarrow \mu^+ \mu^-) &= 1.031(54) \times 10^{-10}. \end{aligned} \quad (20)$$

We can also obtain the ratio of branching fractions for  $B_s$  and  $B_d$  [66]

$$\frac{\text{Br}(B_s \rightarrow \mu^+ \mu^-)}{\text{Br}(B_d \rightarrow \mu^+ \mu^-)} = \frac{\tau_{B_s^H} \hat{B}_{B_d}^{(1)} \Delta M_s}{\tau_{B_d^H} \hat{B}_{B_s}^{(1)} \Delta M_d}. \quad (21)$$

Here we have dropped the terms in  $m_\mu^2/M_{B_q}^2$  since they are negligible. There is a lot of cancellation in this ratio, including of systematic errors in the ratio of bag parameters [see our results in Eq. (12)]. We obtain the result

$$\frac{\text{Br}(B_d \rightarrow \mu^+ \mu^-)}{\text{Br}(B_s \rightarrow \mu^+ \mu^-)} = 0.02706(70). \quad (22)$$

Figure 10 shows our predictions in the SM for the branching fractions from Eqs. (20) and (22) as the grey

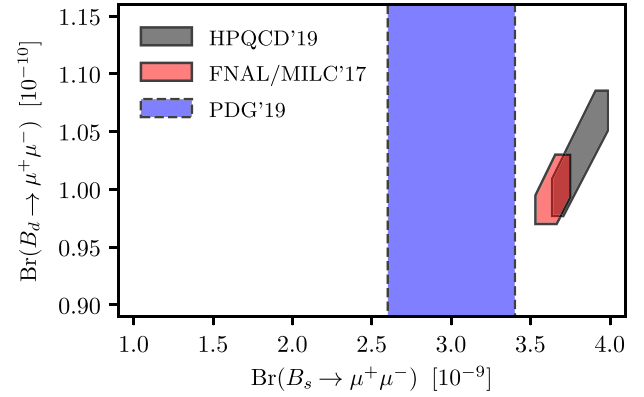


FIG. 10. A comparison of the SM branching fractions for  $B_s$  and  $B_d$  to decay to  $\mu^+ \mu^-$  from lattice QCD results with the current experimental measurements. The grey lozenge shows results from our calculation here of the bag parameters and a ratio to experimental results for  $\Delta M_q$  [Eqs. (20) and (22)]. The red lozenge shows results from a Fermilab/MILC calculation of  $B_d$  and  $B_s$  decay constants, combined with input CKM elements [13]. The blue band shows the current experimental average for  $\text{Br}(B_s \rightarrow \mu^+ \mu^-)$  [11]; only an upper bound exists for  $\text{Br}(B_d \rightarrow \mu^+ \mu^-)$ .

lozenge. The red lozenge shows lattice QCD predictions [13] for the branching fractions using the direct approach where the hadronic parameter needed is the decay constant and this is combined with input for the CKM elements  $V_{tq}$  and  $V_{tb}$ , along with other factors. The errors in the results from [13] are dominated by uncertainties in the CKM elements, which are taken from a global unitarity triangle fit that includes both tree and loop-induced processes.<sup>2</sup>

Figure 10 shows good agreement between the two lattice QCD predictions. This reflects the fact that, as described in Sec. IV D our results for the bag parameters yield CKM elements  $|V_{ts}|$  and  $|V_{td}|$  in agreement with CKM unitarity determinations. Our results imply consistency of the CKM matrix (within uncertainties), and hence the two approaches of using the decay constants plus CKM elements or using the bag parameters and  $\Delta M_q$  will agree.

Note that our results in Eq. (20) include uncertainties in the parameters of Eq. (21). They do not include uncertainties from electromagnetic corrections to the decay process. These are estimated to lead to a reduction of 0.3%–1.1% in the muonic branching fractions in [70]. This is not significant given the current uncertainties in our SM predictions, but will need to be addressed to reduce uncertainties in the future.

The blue band in Fig. 10 shows the current experimental situation. The decay  $B_d \rightarrow \mu^+ \mu^-$  has only been seen with  $3\sigma$  significance [71]. Recent LHCb [72] and ATLAS [73] results give upper bounds to the branching fraction of  $3.4 \times 10^{-10}$  and  $2.1 \times 10^{-10}$ , respectively. These bounds

<sup>2</sup>Note that we include the constraint on the ratio from the July 2019 update of [13].



are outside the range of Fig. 10. For the branching fraction for  $B_s \rightarrow \mu^+ \mu^-$ , the Particle Data Group quotes an average value of  $3.0(4) \times 10^{-9}$  using results from ATLAS [73], CMS [74], and LHCb [72]. The  $\pm 1\sigma$  variation gives the width of the band in Fig. 10.

Although no significant tension between experiment and the SM predictions is visible in this figure, it does give encouragement that we are reaching a point where further reductions in uncertainties will start to give serious SM constraints, at least for  $B_s \rightarrow \mu^+ \mu^-$ .

### F. Contributions to $\Delta\Gamma$

Another physical observable from neutral  $B$  meson systems is that of the decay width difference of the eigenstates,  $\Delta\Gamma$ . This has been measured for the  $B_s$  at 13% of the average width, but only an upper limit exists for the  $B_d$  [11]. The prediction for the width differences in the SM is given in [16,75] in terms of the matrix elements of several 4-quark operators. We give results here for the matrix elements of those labeled  $R_0$ ,  $R_1$ , and  $\tilde{R}_1$  in [75].

$R_0$  [75] is a combination of  $O_1$ ,  $O_2$ , and  $O_3$  which is a  $1/m_b$ -suppressed operator up to corrections of  $\mathcal{O}(\alpha_s^2)$  times a leading order operator,

$$R_0 = O_2 + O_3 + \frac{O_1}{2} + \alpha_s(0.345O_1 + 0.637O_3), \quad (23)$$

evaluating the radiative corrections at  $\mu = m_b$ . We can obtain the matrix elements for this operator through  $\mathcal{O}(\alpha_s)$  from our lattice calculation.  $R_1$  and  $\tilde{R}_1$  are proportional to  $\mathcal{O}_4$  and  $\mathcal{O}_5$ , respectively, and will be discussed further below.

The matrix elements for  $R_0$  in Eq. (23) can be rewritten in terms of our bag parameters using the definition in Eq. (4):

$$\begin{aligned} \langle B_q | R_0 | B_q \rangle = & -f_{B_q}^2 M_{B_q}^2 \left( \frac{M_{B_q}}{m_b(\mu) + m_q(\mu)} \right)^2 \\ & \times \left[ \frac{5}{3} B_{B_q}^{(2)}(\mu) - \frac{1}{3} B_{B_q}^{(3)}(\mu) (1 + 0.637\alpha_s) \right. \\ & \left. - \frac{4}{3} B_{B_q}^{(1)}(\mu) (1 + 0.690\alpha_s) \left( \frac{m_b(\mu) + m_q(\mu)}{M_{B_q}} \right)^2 \right]. \end{aligned} \quad (24)$$

Writing it in this way makes clear (setting  $B_{B_q}^{(n)}$  to 1 and  $\alpha_s$  to zero) the expected cancellation at leading order to leave matrix elements that are  $\mathcal{O}(1/m_b)$ . Our evaluation of the term in square brackets above yields

$$\begin{aligned} \langle B_d | R_0 | B_d \rangle &= -f_{B_d}^2 M_{B_d}^2 (3\eta_3^d) \times 0.22(12), \\ \langle B_s | R_0 | B_s \rangle &= -f_{B_s}^2 M_{B_s}^2 (3\eta_3^s) \times 0.27(11). \end{aligned} \quad (25)$$

Note that the uncertainties here include both those for missing  $\alpha_s^2$  terms in the matching of lattice QCD operators to continuum operators and also the effect of missing  $\alpha_s^2$  terms in the definition of  $R_0$ . This latter uncertainty is estimated by calculating the size of the  $\alpha_s$  corrections in Eq. (23) and multiplying by  $\alpha_s$ . This gives a 35% uncertainty, which dominates the error quoted in Eq. (25). To assist with numerical evaluation we have replaced the square ratio of masses in Eq. (24) with  $3\eta_3^q$ ; values for this can be found in Table I. Equation (25) avoids the use of a perhaps somewhat arbitrary definition of a bag parameter for  $R_0$  given in [75]. The numerical factors show clearly that this is a  $1/m_b$ -suppressed operator by being of size  $\Lambda_{\text{QCD}}/m_b \approx 10\%$ .

$R_1$  and  $\tilde{R}_1$  are defined as [75]

$$R_1^q = \frac{m_q}{m_b} O_4, \quad \tilde{R}_1^q = \frac{m_q}{m_b} O_5. \quad (26)$$

The matrix elements for  $B_s$  and  $B_d$  can then be determined from our results in Table V. Our bag parameters for  $O_4$  and  $O_5$  are given in Table VI. In [75] bag parameters for  $R_1$  and  $\tilde{R}_1$  are defined in such a way as to set the squared mass ratios in Eq. (4) to 1. This means that the bag parameters for  $R_1$  and  $\tilde{R}_1$  for the definition in [75] can be recovered from our bag parameters by multiplying by  $3\eta_4^q/7$  and  $3\eta_5^q/5$ , respectively. These factors are larger than 1. Note, however, that the impact of both  $R_1$  and  $\tilde{R}_1$  on  $\Delta\Gamma$  is tiny because of the  $m_q/m_b$  factors in their definition.

Matrix elements of the numerically more important  $R_2$  and  $\tilde{R}_2$  operators [75], along with those of  $R_3$  and  $\tilde{R}_3$ , cannot be directly obtained from our current results because they contain derivatives on the light quark fields inside the 4-quark operator. Results of the calculations of these matrix elements are discussed separately in [76].

### V. CONCLUSIONS

We give results from the first “second-generation” lattice QCD calculation of the matrix elements that contribute to  $B_s$  and  $B_d$  mixing in and beyond the Standard Model. We include  $c$  quarks in the sea for the first time and have a range of  $u/d$  quark masses (taken to be equal) that go down to the physical value. We use radiatively improved NRQCD for the  $b$  quark action. By calculating the ratio of the matrix elements of the 4-quark operators to the square of the decay constant times mass (proportional to a quantity known as the bag parameter) we obtain results with very little dependence on the  $u/d$  quark mass or the lattice spacing. This gives us more accurate results than previous calculations for these ratios, and the associated bag parameters, for all five  $\Delta B = 2$  operators.

Our key results are given in Tables V and VI. Table V gives the ratio of matrix elements to  $(fM)^2$  with our final physical values given in the last row. These are the numbers

that should be used to reconstruct the 4-quark operator matrix elements by multiplying by  $(f_{B_q} M_{B_q})^2$ . Table VI converts these ratios into bag parameters, defined in Eq. (4). These numbers can be compared to unity, the result expected in the vacuum saturation approximation. Our error budget for the bag parameters is given in Table VII. We have uncertainties of 4%–7% for the individual bag parameters. This uncertainty is dominated by missing higher orders in the perturbative matching to the continuum 4-quark operators. The uncertainty is reduced to around 2% in the ratio of bag parameters for  $B_s$  to  $B_d$ , since this renormalization cancels. The correlations between results for different operators are given in Table X of Appendix C. Our  $B_s$  to  $B_d$  ratios are now accurate enough to see that they are above 1 for  $O_1$  and  $O_3$  and below 1 for  $O_4$  and  $O_5$ .

Our results for the key  $O_1$  bag parameters that appear in SM phenomenology are [repeating Eq. (12)]

$$\begin{aligned}\hat{B}_{B_s}^{(1)} &= 1.232(53), \\ \hat{B}_{B_d}^{(1)} &= 1.222(61), \\ \frac{\hat{B}_{B_s}^{(1)}}{\hat{B}_{B_d}^{(1)}} &= 1.008(25),\end{aligned}\quad (27)$$

where we give the RGI bag parameter as defined in Eq. (11). Multiplying by decay constant values obtained using HPQCD's approach to  $b$  physics with HISQ quarks [40] by the Fermilab/MILC Collaboration [13], we also obtain [repeating Eq. (13)]

$$\begin{aligned}f_{B_s} \sqrt{\hat{B}_{B_s}^{(1)}} &= 0.2561(57) \text{ GeV}, \\ f_{B_d} \sqrt{\hat{B}_{B_d}^{(1)}} &= 0.2106(55) \text{ GeV}, \\ \xi &= 1.216(16).\end{aligned}\quad (28)$$

In Sec. IV we discuss the phenomenology from our results. We obtain values for  $\Delta M$  for  $B_s$  and  $B_d$  in Eq. (17) to be compared to experiment.

Alternatively, and more usefully, we can combine our results with experiment to obtain the CKM elements  $|V_{ts}|$  and  $|V_{td}|$  and their ratio. These values are given in Eq. (18), and Fig. 9 shows the constraints they give in the  $V_{td} - V_{ts}$  plane. Our results are the most accurate determinations of these CKM elements using lattice QCD and show good consistency with determinations from tree-level processes assuming CKM unitarity. This means that we see no signs of new physics in neutral  $B$ -meson oscillations at this improved level of accuracy.

We derive results, by taking a ratio to  $\Delta M$ , for the branching fractions for  $B_s$  and  $B_d$  to decay to  $\mu^+ \mu^-$ , a key mode for new physics searches at LHC. Our results are given in Eqs. (20) and (22). Figure 10 shows a comparison of our predictions to those from the recent Fermilab/MILC

calculation of decay constants [13] along with the current experimental picture. This is encouraging for future tests of new physics contributions to these rare decay processes.

Finally, we give values in Eq. (25) and below Eq. (26) for the matrix elements for the  $R_0$ ,  $R_1$ , and  $\tilde{R}_1$  operators that contribute to the SM prediction for the width difference  $\Delta\Gamma$ .

To improve accuracy further in the future requires improving the matching of the lattice QCD 4-quark operators to those in the continuum; this is the dominant source of uncertainty in our bag parameters. Since lattice QCD perturbation theory is so hard, a renormalization method that can be implemented within the lattice calculation and then matched perturbatively to  $\overline{\text{MS}}$  in the continuum could be preferable. A symmetric momentum-subtraction scheme (RI-SMOM) has been found to work well for kaon mixing calculations [37] but attention must be paid to removing nonperturbative artifacts in these schemes if high accuracy is to be achieved [77]. Such a method would need to be implemented with a relativistic quark action on lattices with fine enough lattice spacing to allow quark masses close to that of the  $b$  for  $am \lesssim 1$ . This has been a very successful strategy for HISQ quarks for  $B$ -meson decay constants [13,40,41], but in that case the decay constants calculated with HISQ do not need any renormalization. Calculating 4-quark operator matrix elements is much harder, but still feasible. The ETM work with twisted mass quarks [7] is encouraging for this program (although they used  $n_f = 2$  gluon fields and RI-MOM renormalization) as is the work by RBC/UKQCD using domain-wall quarks [9] (although they have only calculated  $B_s$  to  $B_d$  ratios so far). It seems clear that in the next few years improvements in this direction will be possible, pushing uncertainties on bag parameters down to the  $\sim 2\%$  level. This will allow  $|V_{ts}|$  and  $|V_{td}|$  to be determined to 1%.

## ACKNOWLEDGMENTS

We are grateful to the MILC Collaboration for the use of their gauge configurations and code. The results described here were obtained using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service as part of STFC's DiRAC facility. This work was funded by STFC, the Royal Society, the Wolfson Foundation, and the U.S. DOE and National Science Foundation. We thank Chris Bouchard, Elvira Gámiz, and Alex Lenz for useful discussions. One of the authors (G.P.L.) is grateful to the Department of Applied Mathematics and Theoretical Physics, Cambridge University, for their hospitality during the two visits when much of this analysis was done.

## APPENDIX A: FITTING PROTOCOLS

We extract mixing amplitudes and decay constants by fitting Monte Carlo data for 2-point and 3-point matrix correlators for each meson:

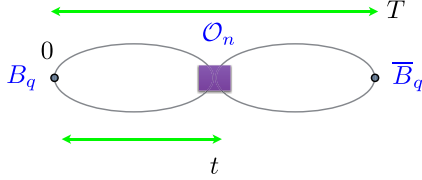


FIG. 11. A schematic diagram of the 3-point function for  $B - \bar{B}$  mixing.  $\mathcal{O}_n$  marks the insertion at time  $t$  of a 4-quark operator, when the meson and antimeson operators are located at times 0 and  $T$ .

$$\mathbf{G}(t) \equiv \sum_{\mathbf{x}} \langle 0 | \mathbf{s}(\mathbf{x}, t) \mathbf{s}^T(0) | 0 \rangle, \quad (\text{A1})$$

$$\mathbf{G}_\beta(t, T) \equiv \sum_{\mathbf{x}, \mathbf{y}} \langle 0 | \mathbf{s}(\mathbf{x}, T) \mathcal{O}_\beta(\mathbf{y}, t) \mathbf{s}^T(0) | 0 \rangle, \quad (\text{A2})$$

where  $\mathbf{s}$  is a 3-vector of meson sources, the sums over spatial  $\mathbf{x}$  and  $\mathbf{y}$  project onto zero 3-momentum, the times satisfy  $0 < t < T$ , and  $\beta = 1, 2, \dots, 5$  labels the mixing operator (Fig. 11). The sources include a local source, corresponding to

$$J_{A_0}^{(0)} = \bar{\Psi}_q \gamma_5 \gamma_0 \Psi_b^{\text{NR}}, \quad (\text{A3})$$

and two smeared sources: see [2] for details. We also examine the vector of correlators

$$\mathbf{G}^{(1)}(t) \equiv \sum_{\mathbf{x}} \langle 0 | J_{A_0}^{(1)}(\mathbf{x}, t) \mathbf{s}^T(0) | 0 \rangle, \quad (\text{A4})$$

where

$$J_{A_0}^{(1)} = -\frac{1}{2m_b} \bar{\Psi}_q \gamma_5 \gamma_0 \boldsymbol{\gamma} \cdot \nabla \Psi_b^{\text{NR}} \quad (\text{A5})$$

is the leading NRQCD correction to  $J_{A_0}^{(0)}$ . We use the corrected current<sup>3</sup> to evaluate the decay constant (in lattice units):

$$(1 + z_{A_0} \alpha_s) \langle 0 | J_{A_0}^{(0)} + J_{A_0}^{(1)} | B_q \rangle = f_{B_q} M_{B_q}, \quad (\text{A6})$$

where coefficients  $z_{A_0}$  depend on the lattice spacing and were calculated in [2] using [17] (see Table II for the values we use here). The values for  $\alpha_s$  (from [15]) are given in Table IV.

The 2-point and 3-point correlators are calculated in the standard way. For 3-point correlators this involves combining propagators from a local source (at  $t$ ) into “open-meson” propagators [8] which are then closed off at 0 and  $T$  (which cover all time slices away from  $t$ ) with local or smeared meson and antimeson operators. We average

<sup>3</sup>Note that the NRQCD-HISQ temporal axial current that we use here is correct through the same order in  $\alpha_s$  and  $\Lambda_{\text{QCD}}/m_b$  as that of our 4-quark operators. In [2] we used a more highly corrected temporal axial current to determine  $f_B$ .

correlators over 16 values of  $t$  for improved statistics. The smearing functions we use are given in [15].

Fits for the  $B_s$  proceed in two steps. First the 2-point correlators  $\mathbf{G}(t)$  and  $\mathbf{G}^{(1)}(t)$  are fit simultaneously. Then the best-fit amplitudes and energies from that fit are used as priors for a simultaneous fit of all the 3-point correlators  $\mathbf{G}_n(t)$ . Having finished the  $B_s$  fits, we follow the same approach with the  $B_d$  correlators but constraining (via the priors) the  $B_d$ 's fit parameters to be within 20% of the corresponding values for the  $B_s$ . We discuss all of these fits in what follows.

## 1. Fitting two-point correlators

We fit the 2-point correlators to a formula of the form (in lattice units)

$$\mathbf{G}^{\text{fit}}(t; \mathbf{p}) = \sum_{n=0}^{N-1} (e^{-E_n t} \mathbf{c}_n \mathbf{c}_n^T - (-1)^t e^{-E_n^o t} \mathbf{c}_n^o \mathbf{c}_n^{oT}), \quad (\text{A7})$$

$$\mathbf{G}^{(1)\text{fit}}(t) = \sum_{n=0}^{N-1} (e^{-E_n t} j_n \mathbf{c}_n^T - (-1)^t e^{-E_n^o t} j_n^o \mathbf{c}_n^{oT}), \quad (\text{A8})$$

where the fit parameters  $\mathbf{p}$  are composed of all  $\mathbf{c}_n$ ,  $\mathbf{c}_n^o$ ,  $E_n$ , and  $E_n^o$ . Here  $\mathbf{c}_n$  and  $\mathbf{c}_n^o$  are 3-component vectors, and  $j_n$  and  $j_n^o$  scalars, where

$$\mathbf{c}_n = \frac{\langle 0 | \mathbf{s} | E_n \rangle}{\sqrt{2M_n}}, \quad \mathbf{c}_n^o = \frac{\langle 0 | \mathbf{s} | E_n^o \rangle}{\sqrt{2M_n^o}}, \quad (\text{A9})$$

$$j_n = \frac{\langle 0 | J_{A_0}^{(1)} | E_n \rangle}{\sqrt{2M_n}}, \quad j_n^o = \frac{\langle 0 | J_{A_0}^{(1)} | E_n^o \rangle}{\sqrt{2M_n^o}}. \quad (\text{A10})$$

In the exponents,  $E_n$  and  $E_n^o$  are the energies of the lowest-lying states with zero 3-momentum that couple to the sources. The second (oscillating in time) term in each correlator is due to taste doubling caused by the staggered-quark HISQ action for the light quarks (see [33]).  $M_n$  and  $M_n^o$  are the physical masses corresponding to states  $|E_n\rangle$  and  $|E_n^o\rangle$ , respectively. We keep  $N = 6$  terms, but fit results are the same for any  $N \geq 5$ .

We use a Bayesian fit procedure [78]. Fits for the  $B_s$  mesons on our coarsest lattices (0.15 fm) use the following Bayesian priors for the energies,

$$\begin{aligned} \log(E_0) &= \log(0.6(3)), & \log(\Delta E_n) &= \log(0.50(25)), \\ \log(E_0^o) &= \log(0.90(45)), & \log(\Delta E_n^o) &= \log(0.50(25)), \end{aligned} \quad (\text{A11})$$

where  $\Delta E_n \equiv E_n - E_{n-1}$ , and the logarithms indicate log-normal priors for energies. Energies are rescaled in proportion to the lattice spacing to make priors for the other lattices. The priors for local and smeared amplitudes are

$$\begin{aligned}
\log(\mathbf{c}_n(\text{loc.})) &= \log(0.2(8)), & \mathbf{c}_n(\text{smeared}) &= 1(4), \\
\log(\mathbf{c}_n^o(\text{loc.})) &= \log(0.2(8)), & \mathbf{c}_n^o(\text{smeared}) &= 1(4),
\end{aligned}
\tag{A12}$$

on the coarsest lattices. Amplitudes for the local source are rescaled by  $a^{3/2}$  for the other lattices. The smeared sources are designed to be lattice-spacing independent, and so we use the same prior for the other lattices. Finally the priors for parameters  $j_n$  and  $j_n^o$  are

$$j_n = -0.015(60), \quad j_n^o = 0.02(8) \tag{A13}$$

on the coarse lattice, and again scale as  $a^{3/2}$  for other lattices.

These central values for these priors were based upon fit values for the ground state  $B_s$ . The uncertainties assigned to the priors are large: for example, the priors are typically 2000–5000 times broader than the final fit errors for the ground state parameters that we need for our analysis. Replacing the central values by random values drawn from the prior distributions leaves our results unchanged within errors.

We need to apply SVD cuts to the data's correlation matrix because of the large number of correlators being fit. The procedure for determining the SVD cuts is described in Appendix D; typically the cuts affect less than half of the data modes. Fits for  $\mathbf{G}(t)$  were constrained to  $t$  values between the values for  $t_{\min}$  and  $t_{\max}$  shown in Table VIII; data from larger  $t$ 's was too noisy to be useful. To keep the number of fit points down (see Appendix D), we restricted the fits for  $G^{(1)}(t)$  to the range of  $t$ 's between  $(t_{\min} + t_{\max})/2$  and  $t_{\max}$ .

## 2. Fitting three-point correlators

The fit function for the 3-point correlators is substantially more complicated:

$$\begin{aligned}
\mathbf{G}_\beta^{\text{fit}}(t, T; \mathbf{p}) &= \sum_{n,m=0}^{N-1} e^{-E_n t} \mathbf{c}_n V_{nm}(O_\beta) \mathbf{c}_m^T e^{-E_m(T-t)} \\
&\quad - (-1)^{T-t} \sum_{n,m=0}^{N-1} e^{-E_n t} \mathbf{c}_n V_{nm}^o(O_\beta) \mathbf{c}_m^{oT} e^{-E_m^o(T-t)} \\
&\quad - (-1)^t \sum_{n,m=0}^{N-1} e^{-E_n^o t} \mathbf{c}_n^o V_{mn}^o(O_\beta) \mathbf{c}_m^T e^{-E_m(T-t)} \\
&\quad + (-1)^T \sum_{n,m=0}^{N-1} e^{-E_n^o t} \mathbf{c}_n^o V_{nm}^{oo}(O_\beta) \mathbf{c}_m^{oT} e^{-E_m^o(T-t)},
\end{aligned}
\tag{A14}$$

where the fit parameters  $\mathbf{p}$  include all of the 2-point correlator parameters plus the  $V_{nm}$ ,  $V_{nm}^o$ , and  $V_{nm}^{oo}$ .

We fit the 3-point amplitudes over the range  $t_{\min} \leq t \leq T - t_{\min}$  for the values of  $T$  shown in Table VIII. Parameters  $\mathbf{c}_n$ ,  $\mathbf{c}_n^o$ ,  $E_n$ , and  $E_n^o$  are the same as in the 2-point correlators; we use the results from the fits to the 2-point correlators as priors for these parameters in our 3-point fits. The priors on the coarsest lattices for each of the mixing amplitudes  $V(\beta)$ ,  $V^o(\beta)$ , and  $V^{oo}(\beta)$  are

$$\begin{aligned}
V(O_1) &= 0.03(12), & V(O_2) &= -0.03(12), \\
V(O_3) &= 0.005(20), & V(O_4) &= 0.04(16), \\
V(O_5) &= 0.025(100);
\end{aligned}
\tag{A15}$$

TABLE VIII. Simulation results for  $\langle B_q | O_n | \bar{B}_q \rangle_{\text{latt}} / (f_{B_q} M_{B_q})^2$  for  $B_q = B_s, B_d$  mesons. Results are presented for each of the configuration data sets described in Table III. Fit ranges for 2-point ( $t_{\min} \leq t \leq t_{\max}$ ) and 3-point ( $t_{\min} \leq t \leq T - t_{\min}$ ) correlators are tabulated. Sample  $\chi^2$ 's per degree of freedom from fits to both sets of correlators, with SVD and prior noise (see Appendix D 4), are also listed.

Meson	Set	$t_{\min}$	$t_{\max}$	$T$	$\langle O_1 \rangle / (fM)^2$	$\langle O_2 \rangle / (fM)^2$	$\langle O_3 \rangle / (fM)^2$	$\langle O_4 \rangle / (fM)^2$	$\langle O_5 \rangle / (fM)^2$	$\chi^2/\text{dof}$ [dof]
$B_s$	1	4	17	8–12	2.274(29)	−1.887(24)	0.3646(68)	3.041(35)	1.870(23)	0.97 [258]
$B_s$	2	4	17	8–12	2.336(29)	−1.939(21)	0.3625(75)	3.199(33)	1.978(22)	1.11 [258]
$B_s$	3	4	17	8–12	2.315(31)	−1.923(26)	0.3638(73)	3.162(39)	1.951(24)	1.01 [258]
$B_s$	4	4	22	10–14, 17	2.367(27)	−1.979(21)	0.3720(70)	3.382(34)	2.060(22)	0.92 [324]
$B_s$	5	4	22	10–14, 17	2.319(26)	−1.955(20)	0.3636(70)	3.288(34)	1.994(22)	1.13 [324]
$B_s$	6	4	22	10–14, 17	2.365(32)	−1.999(25)	0.3747(80)	3.366(40)	2.037(25)	0.99 [324]
$B_s$	7	5	33	12–16, 19	2.312(32)	−2.029(29)	0.3835(83)	3.548(47)	2.090(28)	1.03 [399]
$B_d$	1	4	17	8–12	2.238(56)	−1.869(39)	0.362(14)	3.054(59)	1.861(41)	1.00 [258]
$B_d$	2	4	17	8–12	2.283(70)	−1.875(43)	0.345(19)	3.232(69)	2.016(49)	1.05 [258]
$B_d$	3	4	17	8–12	2.228(74)	−1.854(49)	0.324(19)	3.232(77)	1.955(53)	1.07 [258]
$B_d$	4	4	22	10–14, 17	2.377(56)	−1.941(38)	0.333(14)	3.539(67)	2.158(44)	1.06 [324]
$B_d$	5	4	22	10–14, 17	2.332(59)	−1.870(38)	0.328(15)	3.341(61)	2.011(43)	0.99 [324]
$B_d$	6	4	22	10–14, 17	2.265(97)	−1.965(67)	0.379(25)	3.49(10)	2.086(71)	1.08 [324]
$B_d$	7	5	33	12–16, 19	2.332(70)	−1.984(68)	0.356(18)	3.68(10)	2.153(63)	1.08 [399]



these are scaled in proportion to  $a^3$  for the other lattices. Note that  $V_{n,m}(\beta)$  and  $V_{n,m}^{oo}(\beta)$  are symmetric under the interchange of  $n$  and  $m$ . We are interested in the ground-state value for

$$V_{00}(O_\beta) = \frac{\langle E_0 | O_\beta | E_0 \rangle}{2M_0}. \quad (\text{A16})$$

We introduce two simplifications to the analysis that make our fits run 20–100 times faster, without affecting fit results or precision. The first simplification is to replace both the data and the fit function in the fits with their sums over  $t$ ,

$$\begin{aligned} \mathbf{G}_\beta(t, T) &\rightarrow \sum_{t=t_{\min}}^{T-t_{\min}} \mathbf{G}_\beta(t, T), \\ \mathbf{G}_\beta^{\text{fit}}(t, T; \mathbf{p}) &\rightarrow \sum_{t=t_{\min}}^{T-t_{\min}} \mathbf{G}_\beta^{\text{fit}}(t, T; \mathbf{p}), \end{aligned} \quad (\text{A17})$$

while keeping the same fit parameters [79]. This reduces the number of data points to be fit for our 0.09 fm lattice, for example, from 1050 to 180. Note that the Monte Carlo data for  $\mathbf{G}_n(t, T)$  do not vary much with  $t$ , as expected from Eq. (A14).

The second simplification is to marginalize all fit parameters other than those associated with the ground state [80]. We do this by splitting the fit function [Eq. (A14)] into two parts, one that involves only the ground state (i.e., either the  $B_s$  or  $B_d$ ) and the other with the remaining terms:

$$\mathbf{G}_\beta^{\text{fit}} \equiv \mathbf{c}_0 V_{00}(O_\beta) \mathbf{c}_0^T e^{-E_0 T} + \Delta \mathbf{G}_\beta^{\text{fit}}. \quad (\text{A18})$$

We then replace the fit data  $\mathbf{G}_\beta^{\text{lat}}$  by

$$\mathbf{G}_\beta^{\text{lat}}(t, T) \rightarrow \mathbf{G}_\beta^{\text{lat}}(t, T) - \Delta \mathbf{G}_\beta^{\text{fit}}(t, T; \mathbf{p}_{\text{prior}}), \quad (\text{A19})$$

where the prior values for the fit parameters are used in  $\Delta \mathbf{G}_\beta^{\text{fit}}$ . At the same time, we replace the fit function by just its ground-state term:

$$\mathbf{G}_\beta^{\text{fit}}(t, T; \mathbf{p}) \rightarrow \mathbf{c}_0 V_{00}(O_\beta) \mathbf{c}_0^T e^{-E_0 T}. \quad (\text{A20})$$

This reduces the number of fit parameters from 450 to 9 (for  $N = 6$ ). Marginalization works particularly well here because we have excellent priors for the amplitudes and energies, from the 2-point correlators, and because the mixing parameters enter the fit function linearly. Also the marginalized fit function is  $t$  independent, making the first simplification (summing over  $t$ ) quite natural.

Again we need SVD cuts, but the need is greatly reduced by summing over  $t$ . We used the method outlined in

Appendix D; typically the cuts modified around 70% of the data modes.

We tabulate simulation results [using Eqs. (A6), (A9), (A10), and (A16)] for the dimensionless ratio

$$\frac{\langle B_q | O_n | B_q \rangle}{f_{B_q}^2 M_{B_q}^2} \quad (\text{A21})$$

with  $B_q = B_s, B_d$  in Table VIII. This table also shows sample values of  $\chi^2$  from the various (2-point and 3-point) correlator fits when we include random SVD and prior noise, as discussed in Appendix D 4. Without noise, the  $\chi^2$ 's per degree of freedom are much smaller than 1.0, as expected. Also following Appendix D 4, we tested the uncertainties from our fits using simulated data. Fits to simulated data reproduced the input parameters to within errors.

We verified that marginalization and averaging over  $t$  have negligible effect on our results. With our smallest lattice spacing (0.09 fm), for example, undoing both optimizations gives the following values:

$$\frac{\langle B_s | O_n | B_s \rangle}{f_{B_s}^2 M_{B_s}^2} = \begin{cases} 2.307(35) & n = 1, \\ -2.025(31) & n = 2, \\ 0.3847(86) & n = 3, \\ 3.547(51) & n = 4, \\ 2.094(31) & n = 5. \end{cases} \quad (\text{A22})$$

These agree well with the values in Table VIII (for set 7), but took far longer to compute.

In Table IX, we show sample error budgets for these quantities from simulations on the 0.09 fm lattice; others are similar. The dominant source of uncertainty is from the Monte Carlo statistics.

TABLE IX. Sample error budgets from simulations on the 0.09 fm lattice (set 7) for  $\langle B_q | O_n | B_q \rangle / (f_{B_q} M_{B_q})^2$  for  $B_s$  and  $B_d$  mesons. Percentage errors coming from Monte Carlo statistics, the fit priors, and the SVD cuts are shown; these are added in quadrature to give the total error.

$B_s$	$\langle O_1 \rangle / (fM)^2$	$\langle O_2 \rangle / (fM)^2$	$\langle O_3 \rangle / (fM)^2$
Statistics	1.25	1.15	2.00
Prior	0.28	0.31	0.47
SVD	0.42	0.74	0.56
Total	1.35	1.41	2.13
$B_d$	$\langle O_1 \rangle / (fM)^2$	$\langle O_2 \rangle / (fM)^2$	$\langle O_3 \rangle / (fM)^2$
Statistics	2.59	2.13	4.25
Prior	0.89	1.00	1.87
SVD	1.38	2.53	1.75
Total	3.07	3.46	4.96

### 3. Blind analysis

This analysis was blinded by multiplying the 3-point correlators by a random normalization factor. The random factor was removed only after the entire analysis was completed and this paper written.

## APPENDIX B: CHIRAL FIT

Although we have results at physical pion masses we do not rely on these simply for our final value. We include also results at heavier-than-physical pion masses, which are statistically more precise, by using a fit to the dependence on the pion mass based on chiral perturbation theory. This gives the coefficients of the nonanalytic “chiral logarithms” in  $m_\pi^2 \log(m_\pi^2)$ ; in addition we include analytic terms to allow both for staggered quark discretization effects, the unphysically heavy  $u/d$  quark masses in the sea, and for mistuning of valence quark masses. Performing a fit to results at multiple pion masses then tests the dependence expected from chiral perturbation theory.

Our principal results consist of values for the “reduced” matrix elements of the 4-quark operators,

$$R_q^n \equiv \frac{\langle B_q | O_n | \bar{B}_q \rangle_{\overline{\text{MS}}}^{(m_b)}}{(f_{B_q} M_{B_q})^2}, \quad (\text{B1})$$

on each configuration set (Table V). We fit the  $R_q^n$  to the form

$$\begin{aligned} R(m_l, m_s) = & R(m_l^{\text{phys}}, m_s^{\text{phys}})(1 + p_{\log} \chi_{\log} + p_J 3g^2 J \\ & + p_{a^4} \delta X_{a^4} + p_{m_\pi^2 a^2} \delta X_{m_\pi^2 a^2} p_l \delta x_l \\ & + p_{l2} (\delta x_l)^2 + p_v \delta v), \end{aligned} \quad (\text{B2})$$

where we suppress the indices  $n$  and  $q$  on each term and each parameter,  $p$ , for clarity. The functions used in each term are discussed below.  $\chi_{\log}$  is given in Eq. (B6);  $J$  in Eq. (B7);  $\delta X$  in Eq. (B9);  $\delta x_l$  in Eq. (B11); and  $\delta v$  in Eq. (B12). Note that this fit is done after applying the additional uncertainties discussed in Sec. III B to allow for matching and discretization effects.

To derive this form, we make use of the results in [81], where Appendix A gives the dependence on light meson masses of the bag parameters at one loop in heavy meson staggered chiral perturbation theory. This builds on the continuum heavy meson chiral perturbation theory results of [82]. There is a lot of cancellation of chiral logarithms between 4-quark operator matrix elements and decay constants so that, as we discuss below, the remaining chiral logarithm terms [ $\chi_{\log}$  and  $J$  in Eq. (B2)] in the bag parameters (and equivalently in  $R$ ) have small coefficients. This expected very benign dependence on the light quark mass is another reason for working with the bag

parameters as we do here, rather than the 4-quark operator matrix elements.

The chiral perturbation theory for the bag parameters is given in [81] in the form (using our notation)

$$R_q^n = \beta_n \left( 1 \pm S^q + \tilde{T}_n^q + \frac{\beta'_n}{\beta_n} (Q_n^q + \tilde{Q}_n^q) \right). \quad (\text{B3})$$

Here  $\beta_n$  is the low-energy constant (value at zero pion mass) for  $R^n$ , and  $\beta'_n$  is the equivalent term for 4-quark matrix elements between vector heavy-light mesons. For  $O_1$ ,  $\beta'_1 = \beta_1$ .  $S$  comes from “tadpole” diagrams (with + for  $n = 1, 2$ , and 3 and – for  $n = 4, 5$ ) and  $Q$  from “sunset” diagrams that connect pseudoscalar and vector mesons.  $\tilde{T}$  and  $\tilde{Q}$  are “wrong-spin” tadpole and sunset terms, respectively. Below we discuss the content of these functions in terms of the important nonanalytic chiral logarithms and the effect on these of the discretization effects in the staggered quark formalism. This enables us to transcribe Eq. (B3) into the simpler Eq. (B2) that we will use. We now discuss each of these terms in turn.

### 1. Tadpole diagrams

The results in [81] are given in terms of meson masses that include those for pions of different tastes that appear in the staggered quark formalism. Thus a term that would be a simple chiral logarithm in the continuum can appear in a number of guises, one of which is as an average over the masses of all tastes of pion. On fine enough lattices this will become a continuum logarithm plus discretization effects. In fact, for the fully unquenched case that we study here ( $m^{\text{val}} = m^{\text{sea}}$ ), staggered chiral perturbation theory typically arranges itself to cancel taste effects inside chiral logarithms so that nonanalyticities in  $a^2$  cancel as  $m_{u/d} \rightarrow 0$  (see Appendix A of [83]). This also happens here. The chiral logarithm in  $R_d$  from tadpole diagrams [ $S$  in Eq. (B3)] appears in the form

$$-\frac{1}{16} \sum_{\text{tastes}, t} \frac{m_{\pi,t}^2}{\Lambda_\chi^2} \log \frac{m_{\pi,t}^2}{\mu_\chi^2} + \frac{1}{2} \frac{m_{\pi,I}^2}{\Lambda_\chi^2} \log \frac{m_{\pi,I}^2}{\mu_\chi^2}. \quad (\text{B4})$$

Here  $I$  denotes the singlet (largest mass) pion taste. We can compare this function to the corresponding continuum chiral logarithm

$$-\frac{1}{2} \frac{m_{\pi,P}^2}{\Lambda_\chi^2} \log \frac{m_{\pi,P}^2}{\mu_\chi^2}, \quad (\text{B5})$$

where  $P$  denotes the lightest (Goldstone) pion taste. This comparison is shown in Fig. 12 using taste splittings for HISQ pions for the lattice spacing values that we use in this calculation. We give results for our range of  $m_\pi^2 \equiv m_{\pi,P}^2$  values from the physical point, 0.018 GeV<sup>2</sup>, to 0.09 GeV<sup>2</sup>.  $\Lambda_\chi = 4\pi f_\pi = 1.64$  GeV, and we take  $\mu_\chi = 1.0$  GeV.

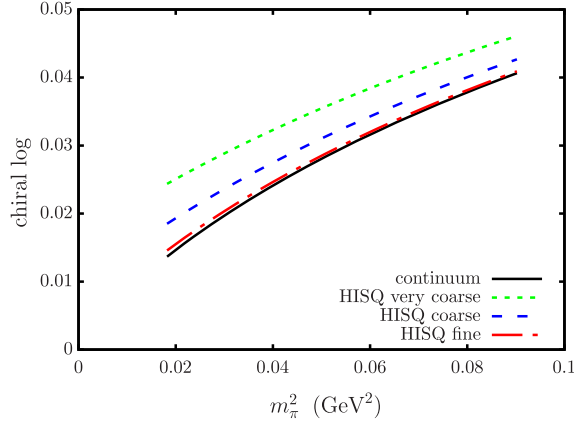


FIG. 12. A comparison of chiral logarithm terms in  $m_\pi$  that appear in the continuum chiral perturbation theory for  $R_d$  [Eq. (B5)] with those in staggered chiral perturbation theory [Eq. (B4)] for HISQ quarks. The solid black line gives the continuum form [Eq. (B5)], and the dashed blue and red curves give the staggered form in Eq. (B4) for HISQ quarks on very coarse, coarse, and fine lattices, respectively.

The difference between the HISQ and continuum chiral logarithm terms is sufficiently small that we simply allow for that discrepancy in our treatment of discretization effects. This is included through the  $\delta X$  terms in Eq. (B2) discussed below.

We therefore take the chiral logarithm terms in Eq. (B2) with continuum form:

$$\begin{aligned}\chi_{\log,n}^d &= -\frac{1}{2} \frac{m_\pi^2}{\Lambda_\chi^2} \log\left(\frac{m_\pi^2}{\mu_\chi^2}\right) - \frac{1}{6} \frac{m_\eta^2}{\Lambda_\chi^2} \log\left(\frac{m_\eta^2}{\mu_\chi^2}\right) - (\text{phys}), \\ \chi_{\log,n}^s &= -\frac{2}{3} \frac{m_\eta^2}{\Lambda_\chi^2} \log\left(\frac{m_\eta^2}{\mu_\chi^2}\right) - (\text{phys}).\end{aligned}\quad (\text{B6})$$

Here  $m_\eta^2 = (2m_{\eta_s}^2 + m_\pi^2)/3$ , and we use masses of Goldstone taste  $\pi$  and  $\eta_s$  in this expression. “(phys)” denotes the value of the previous expression evaluated for physical masses so that the total right-hand side vanishes at that point.  $m_\eta^2$  changes very little as  $m_\pi$  changes, so these terms in the fit do very little. The parameters  $p_{\log,n}^q$  are given priors  $+1.0(3)$  for  $n = 1, 2, 3$  and  $-1.0(3)$  for  $n = 4, 5$  as this logarithm appears with the opposite sign for  $\mathcal{O}_{4,5}$ . The prior width allows for modification of the coefficients from missing higher order terms in chiral perturbation theory.

## 2. Sunset diagrams

We now turn to the term denoted  $J$  in Eq. (B2) that comes from the sunset diagram term  $Q$  in Eq. (B3). This would also take the form of a chiral logarithm  $m_\pi^2 \log(m_\pi^2/\mu_\chi^2)$  in the infinite heavy meson mass limit in the continuum. For finite heavy meson mass, however,

$J$  is modified by terms that depend on heavy meson mass differences, because there is a pseudoscalar to vector heavy meson transition inside the sunset diagram. The form of  $J$  as a function of pion mass and heavy meson mass difference,  $\Delta$ , is given in Eq. (6.17) of [84], which considers chiral perturbation theory terms for the heavy-light meson decay constant. In that case the appropriate value for  $\Delta$  can include heavy-strange to heavy-light mass differences as well as vector to pseudoscalar mass differences. Here, when we consider  $R_d$  and  $R_s$ , that does not happen and we only have to consider the case where  $\Delta = M_{B(s)^*} - M_{B(s)}$ . Then  $\Delta$  takes the value 45 MeV for  $B^* - B$  [11], and we take the same value for  $B_s^* - B_s$  since any differences are expected [33] and seen to be [11] small. Figure 13 compares the function  $J(m_\pi, \Delta)/\Lambda_\chi^2$  with that of the chiral logarithm to which it is equal when  $\Delta = 0$ . Even though  $\Delta$  is small, and much smaller than  $m_\pi$  through most of the range in which we work, we see  $\Delta$  does have an impact, so that  $J$  has smaller magnitude and gradient in  $m_\pi^2$  than its associated chiral logarithm.

$J$  in Eq. (B2) then takes the form

$$\begin{aligned}J_n^d &= \frac{1}{2} \frac{J(m_\pi, \Delta)}{\Lambda_\chi^2} + \frac{1}{6} \frac{J(m_\eta, \Delta)}{\Lambda_\chi^2} - (\text{phys}), \\ J_n^s &= \frac{2}{3} \frac{J(m_\eta, \Delta)}{\Lambda_\chi^2} - (\text{phys}),\end{aligned}\quad (\text{B7})$$

where  $J(m, \Delta)$  is given in [84].  $J$  is multiplied by  $3g^2$ , where  $g$  is the  $BB^*\pi$  coupling. We take the value of  $g$  as 0.5, based on recent lattice QCD calculations [85–87]. The uncertainty on  $g$ , both from the lattice calculations and also from the effect of missing higher order terms in chiral perturbation theory, is absorbed into the coefficient  $p_J$ .  $p_J$  is the ratio of low-energy constants associated with the bag parameters for vector heavy-light mesons to that for pseudoscalars. For  $\mathcal{O}_1$  we know that this ratio is 1 [88].

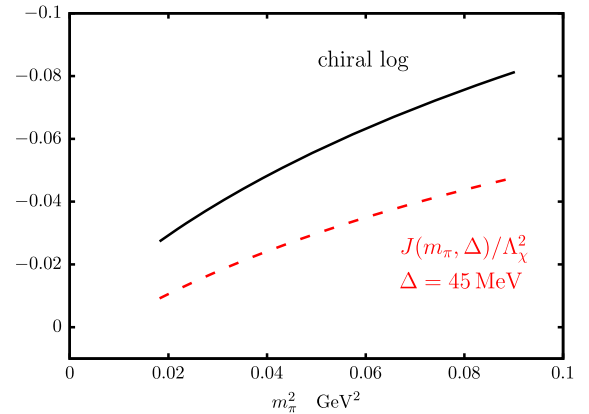


FIG. 13. A comparison of the function  $J(m_\pi, \Delta)/\Lambda_\chi^2$  to the chiral logarithm to which it is equal at  $\Delta = 0$  ( $(m_\pi^2/\Lambda_\chi^2) \log(m_\pi^2/\mu_\chi^2)$ ) as a function of  $m_\pi$ .

For the other operators,  $n = 2-5$ , we do not. We therefore take a prior value and width on  $p_{J,n}^{(d,s)}$  of  $0(1)$ , allowing either sign. For  $p_{J,1}^{(d,s)}$  we take  $1.0(3)$  to allow for uncertainty in  $g^2$ . Note that in the case where  $\Delta = 0$  the coefficient of the chiral logarithm ( $m_\pi^2/\Lambda_\chi^2 \log(m_\pi^2/\mu_\chi^2)$ ) in the chiral perturbation theory for  $O_1$  is  $(3g^2 - 1)/2$  [89], which is small ( $-0.125$ ) when  $g = 0.5$ .

### 3. Wrong spin and other effects

The heavy meson staggered chiral perturbation theory analysis of [81] showed that  $O_{1,2,3}$  and  $O_{4,5}$  can mix through wrong-spin staggered taste effects [ $\tilde{T}$  and  $\tilde{Q}$  in Eq. (B3)]. The size of these terms depends on the size of the light meson taste splittings for the staggered action. For the asqtad action used for light quarks in [8] they were of some concern, and it was important to include these effects explicitly in a full fit to all five operators. For the HISQ action that we use here these effects are much smaller. The question then becomes whether they are distinct in magnitude or form from discretization effects from other sources that are already included in our analysis.

The wrong-spin contributions from tadpole diagrams for  $B_d$  involve differences of chiral logarithms for different taste pions (hence canceling in the absence of taste-splitting effects) along with hairpin terms that have coefficients  $a^2\delta'_V$  and  $a^2\delta'_A$  that are themselves the size of the unit of taste splitting [83]. As an illustration of the impact of these terms we examine the terms that are differences of chiral logarithms. These appear in three “signatures”:  $V - A$ ,  $P - I$ , and  $-P + 2T - I$ . Here the letter denotes the pion taste, ordered in increasing mass as  $P, A, T, V, I$ . Figure 14 illustrates the size and behavior of these terms for the  $-P + 2T - I$  example that mixes  $O_2$  and  $O_3$ , the simplest because

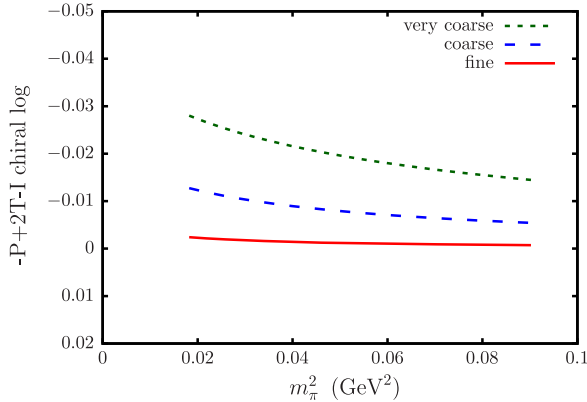


FIG. 14. The function  $\tilde{X}$ , defined in Eq. (B8), plotted for HISQ light quarks against the square of the pion mass. We give three curves, for lattice spacing values corresponding to our very coarse, coarse, and fine lattices. This function appears in the “wrong-sign” tadpole terms in heavy meson staggered chiral perturbation theory.

there are no additional hairpin corrections. The function plotted is

$$\tilde{X} = -\frac{m_{\pi,P}^2}{\Lambda_\chi^2} \log\left(\frac{m_{\pi,P}^2}{\mu_\chi^2}\right) + 2\frac{m_{\pi,T}^2}{\Lambda_\chi^2} \log\left(\frac{m_{\pi,T}^2}{\mu_\chi^2}\right) - \frac{m_{\pi,I}^2}{\Lambda_\chi^2} \log\left(\frac{m_{\pi,I}^2}{\mu_\chi^2}\right). \quad (\text{B8})$$

We see that  $\tilde{X}$  falls rapidly with lattice spacing [approximately as  $(a\Lambda)^4$ ] and has a slope with  $m_\pi^2$  that also falls with lattice spacing [approximately as  $(a\Lambda)^2$ ]. This behavior is generic for terms that arise from taste splittings in this way.

The wrong-spin tadpole terms have a variety of coefficients multiplying them that correspond to ratios of 4-quark operator matrix elements (within the groupings 1,2,3 and 4,5). Most of the coefficients for the wrong-spin tadpole terms appearing in the chiral expansion for  $O_y$  are of the form  $\beta_x/(4\beta_y)$  where  $\beta_x$  is the low energy constant associated with operator  $O_x$ . The exception is  $O_1$ , where the coefficient is  $2(\beta_2 + \beta_3)/\beta_1$ . If all the 4-quark operator matrix elements were of the same size, then the coefficients would be  $\mathcal{O}(1/4)$ .  $O_3$  and  $O_5$  have smaller matrix elements than the others, however, if we consider the vacuum saturation approximation. Hence  $\beta_x/\beta_y$  can be of size 2 for  $O_5$  and 5 for  $O_3$ .

We are already including  $a^2$  and  $a^4$  errors in Eq. (10), but contributions from wrong-sign tadpole terms differ between  $B_d$  and  $B_s$  mesons. The largest contributions are for the  $B_d$  meson; we allow for them and similar terms that arise from sunset diagrams [and so contain differences of  $J(m_{\pi,t}, \Delta)$ ] along with residual right-sign taste effects by including terms

$$\delta X_{a^4}^d = (a\Lambda)^4, \quad \delta X_{m_\pi^2 a^2}^d = \alpha_s (a\Lambda)^2 \delta_{m_\pi^2}, \quad (\text{B9})$$

in the chiral fit, Eq. (B2). We include them with the chiral fit rather than with other discretization effects in Eq. (10) because they arise from the staggered quark action and hence carry no  $am_b$  dependence. Here

$$\delta_{m_\pi^2} \equiv \frac{m_\pi^2 - 0.054}{0.072} \quad (\text{B10})$$

allows for  $m_\pi^2$  dependence; it varies between  $-0.5$  and  $0.5$  over our range of parameters. Similar terms are not needed for  $B_s$  mesons because the effects are smaller and can be simulated by Eq. (10). We take the priors  $p_{a^4}$  and  $p_{m_\pi^2 a^2}$  to be  $0(2)$  since, although this is an unnecessarily broad prior for some  $n$ , it allows a reasonable size for all the possibilities.



#### 4. Analytic terms

The three terms given symbol  $\delta$  on the last line of Eq. (B2) are simple polynomials to account for mistuning of sea and valence quark masses from their physical values. We use

$$\delta x_l = \frac{1}{10} \left( \frac{m_l}{m_s} - \frac{m_l}{m_s} \Big|_{\text{phys}} \right), \quad (\text{B11})$$

where  $m_l$  and  $m_s$  are the sea  $u/d$  and  $s$  quark masses from Table III. The physical value for the  $m_l/m_s$  ratio we take as 27.18(10) from [13]. The factor of 1/10 converts  $m_l/m_s$  to the size of the parameters that appear in chiral perturbation theory as a ratio of meson masses to  $\Lambda_\chi = 4\pi f_\pi$ . Equation (B2) includes terms in  $\delta x_l$  and  $(\delta x_l)^2$ . We take a prior of size 0(1) for each coefficient  $p_l$  and  $p_{l2}$  (for each operator and each  $q$ ). We do not allow for mistuning effects for  $c$  quarks in the sea since we expect these effects to be negligible compared to those from light sea quarks.  $\delta_v$  accounts for the mistuning of light and strange valence masses appropriate to  $B_d$  or  $B_s$ . We take

$$\delta_v^l = \frac{m_\pi^2 - m_{\pi,\text{phys}}^2}{\Lambda_\chi^2}, \quad \delta_v^s = \frac{m_{\eta_s}^2 - m_{\eta_s,\text{phys}}^2}{\Lambda_\chi^2}. \quad (\text{B12})$$

Again the coefficients for this term,  $p_{v,n}^q$ , have prior 0(1) for each  $n$  and  $q$ . We do not include terms for  $b$  quark mass mistuning since the tuning is very accurate and we expect any small mistuning to have a negligible effect on bag parameters.

#### 5. Finite-volume, strong isospin-breaking, and QED effects

Finite-volume effects can be estimated based on chiral perturbation theory. The results in [90] show finite-volume effects of  $\mathcal{O}(1\%)$  for the bag parameters of  $\mathcal{O}_1$  in small lattice volumes of size  $L = 2.5$  fm for physical  $u/d$  quark masses. On our much larger lattices, with a minimum size of  $L = 4.6$  fm at physical  $u/d$  masses for set 3, finite-volume effects will be a lot smaller. We conclude that this is a negligible effect at our current level of uncertainties.

Strong-isospin breaking and electromagnetic effects can also be estimated to be negligible at present. Our bag parameters show very little sensitivity to the  $u/d$  quark mass and our ratios for  $B_s$  to  $B_d$  differ from 1 by at most 10% (Table VI). This suggests that changing  $m_l$  to  $m_d$  should only have a  $\mathcal{O}(0.1\%)$  effect. Effects from the fact that the valence quarks have electromagnetic charge are estimated at below 0.1% for the decay constants in [13]. They come largely from QED effects on the tuning of quark masses. Since bag parameters are less sensitive to both heavy and light quark masses than decay constants, we conclude that QED effects on the bag parameters will be less than 0.1%, and we neglect them. Note that QED effects

can still enter  $\Delta M_q$  or  $\text{Br}(B_q \rightarrow \mu^+ \mu^-)$  through corrections to these processes from adding photons; these effects need to be considered separately.

#### APPENDIX C: CORRELATIONS IN FINAL RESULTS

In this appendix we describe the correlations between the uncertainties in different final results from our analysis. Our principal results consist of values for the reduced matrix elements of the 4-quark operators,  $R_q^n$ , defined in Eq. (B1), evaluated for physical quark masses (Table V). The results for a given meson and different operators are only weakly correlated, as shown in Table X for the  $B_s$  meson. There is more correlation, but still small, between values of the bag parameters [Eq. (2)], because of the normalization factors  $\eta_n^s$  [Eq. (4)]. The means and standard deviations for these quantities are collected in Table XI.

Values of  $R_s^n$  are highly correlated with values of  $R_d^n$ , for the same  $n$ , which is why the ratios  $R_s^n/R_d^n$  have much smaller uncertainties. Uncertainties in these ratios are almost uncorrelated, however, with those in the  $R_q^n$  (correlations are 0.06 or smaller). Thus correlations for the  $B_d$  matrix elements  $R_d^n$  can easily be constructed from the results in Table XI and Table X for  $R_s^n$  and  $R_s^n/R_d^n$ . The ratio

TABLE X. Correlations in the uncertainties of the  $R_s^n$  [Eq. (B1)] for different values of  $n$ . Correlations are also shown for the bag parameters  $B_{B_s}^{(n)}$  [Eq. (2)] and for the ratio  $R_s^n/R_d^n$ . Correlations for  $B_{B_s}^{(n)}/B_{B_d}^{(n)}$  are the same as for  $R_s^n/R_d^n$ .

	$R_s^1$	$R_s^2$	$R_s^3$	$R_s^4$	$R_s^5$
$R_s^1$	1.000	-0.069	0.013	0.041	0.040
$R_s^2$		1.000	-0.039	-0.040	-0.026
$R_s^3$			1.000	0.023	0.012
$R_s^4$				1.000	0.144
$R_s^5$					1.000
	$B_{B_s}^{(1)}$	$B_{B_s}^{(2)}$	$B_{B_s}^{(3)}$	$B_{B_s}^{(4)}$	$B_{B_s}^{(5)}$
$B_{B_s}^{(1)}$	1.000	0.062	0.012	0.037	0.038
$B_{B_s}^{(2)}$		1.000	0.177	0.233	0.150
$B_{B_s}^{(3)}$			1.000	0.170	0.107
$B_{B_s}^{(4)}$				1.000	0.256
$B_{B_s}^{(5)}$					1.000
	$R_s^1/R_d^1$	$R_s^2/R_d^2$	$R_s^3/R_d^3$	$R_s^4/R_d^4$	$R_s^5/R_d^5$
$R_s^1/R_d^1$	1.000	0.296	-0.034	0.064	0.047
$R_s^2/R_d^2$		1.000	0.144	0.068	0.045
$R_s^3/R_d^3$			1.000	0.035	0.018
$R_s^4/R_d^4$				1.000	0.350
$R_s^5/R_d^5$					1.000

TABLE XI. Means and standard deviations for the reduced matrix elements  $R_s^n$  and bag parameters  $B_s^{(n)}$  for each 4-quark operator  $O_n$ , together with values for their  $B_s/B_d$  ratios.

	$R_s^n$	$R_s^n/R_d^n$	$B_s^{(n)}$	$B_s^{(n)}/B_d^{(n)}$
$O_1$	2.1678(928)	1.0081(250)	0.8129(348)	1.0081(250)
$O_2$	-2.1801(1035)	1.0589(242)	0.8169(431)	1.0626(243)
$O_3$	0.4357(288)	1.0886(339)	0.8163(572)	1.0924(340)
$O_4$	3.6532(1480)	0.9558(213)	1.0332(471)	0.9589(214)
$O_5$	1.9448(759)	0.9650(232)	0.9406(384)	0.9668(233)

of bag parameters  $B_{B_s}^{(n)}/B_{B_d}^{(n)}$  is almost equal to  $R_s^n/R_d^n$  (the difference being only from small quark mass effects) and has the same correlation matrix.

## APPENDIX D: SVD CUTS

### 1. The problem

There are three inputs for our least-square fits to sets of correlators: (1) a collection of  $N_s$  random (Monte Carlo) samples  $\mathbf{G}^{(s)}$ , where each sample is packaged as an  $N_G$ -dimensional vector; (2) a (vector) fitting function  $\mathbf{G}(\mathbf{p})$  of fit parameters  $\mathbf{p}$ ; and (3) *a priori* estimates (priors) for the fit parameters.

In the fit, the sample average

$$\bar{\mathbf{G}} \equiv \frac{1}{N_s} \sum_s \mathbf{G}^{(s)} \quad (\text{D1})$$

is assumed to be a random sample drawn from a Gaussian distribution with mean  $\mathbf{G}(\mathbf{p}^*)$  for some set  $\mathbf{p}^*$  of fit parameters, and a covariance matrix given approximately by

$$\mathbf{M}_{\text{cov}} \approx \frac{1}{N_s(N_s - 1)} \sum_s (\mathbf{G}^{(s)} - \bar{\mathbf{G}})(\mathbf{G}^{(s)} - \bar{\mathbf{G}})^T \quad (\text{D2})$$

$$\equiv \mathbf{D} \mathbf{M}_{\text{corr}} \mathbf{D}. \quad (\text{D3})$$

Here  $\mathbf{D}$  is the diagonal matrix of standard deviations,  $\mathbf{D}_{ij} = \delta_{ij} \sigma_{G_i}$ , and  $\mathbf{M}_{\text{corr}}$  is the correlation matrix. The best-fit parameters are obtained by minimizing

$$\chi^2(\mathbf{p}) \equiv \sum_{n=1}^{N_G} \frac{((\bar{\mathbf{G}} - \mathbf{G}(\mathbf{p}))^T \mathbf{D}^{-1} \mathbf{v}_n)^2}{\lambda_n} + \chi_{\text{prior}}^2 \quad (\text{D4})$$

as a function of the parameters  $\mathbf{p}$ , where  $\lambda_n$  and  $\mathbf{v}_n$  are the eigenvalues and eigenvectors of the correlation matrix:

$$\mathbf{M}_{\text{corr}} \mathbf{v}_n = \lambda_n \mathbf{v}_n. \quad (\text{D5})$$

Note that

$$\mathbf{M}_{\text{cov}}^{-1} = \sum_{n=1}^{N_G} \frac{\mathbf{D}^{-1} \mathbf{v}_n \mathbf{v}_n^T \mathbf{D}^{-1}}{\lambda_n}. \quad (\text{D6})$$

$\chi_{\text{prior}}^2(\mathbf{p})$  is the part of  $\chi^2(\mathbf{p})$  associated with the Bayesian priors used in the fit.

The approximation for the covariance matrix, Eq. (D2), causes problems if the number of samples  $N_s$  is insufficiently large compared with the number of data points  $N_G$  [91,92]. In particular, the smaller eigenvalues of the correlation matrix are underestimated. Indeed, it is obvious from Eq. (D2) that there must be  $N_G - N_s$  modes with zero eigenvalue when  $N_s < N_G$ . Underestimating eigenvalues exaggerates their importance in  $\chi^2(\mathbf{p})$  [Eq. (D4)], compromising the fit; and  $\chi^2(\mathbf{p})$  is undefined if there are zero eigenvalues.

The underestimation of small eigenvalues is illustrated in Fig. 15, which shows the ratio of  $\lambda_n^{\text{approx}}/\lambda_n^{\text{exact}}$  for approximate eigenvalues estimated from random samples of different sizes drawn from a simulation of a known distribution (so we know the exact eigenvalues). The small eigenvalues are dragged down to zero by the need for zero modes when  $N_s < N_G$ . They then increase slowly as new samples are added, until  $\lambda_n^{\text{approx}}/\lambda_n^{\text{exact}} \approx 1$  for all  $n$  when  $N_s \gg N_G$ . Note that good approximations for all eigenvalues require  $N_s$  to be 10–100 times larger than  $N_G$ . The figure shows results for  $N_G = 512$  pieces of correlated data; curves for  $N_G = 50$  (or 5000) would be similar, but with more (or less) noise. The range of values covered by the eigenvalues also has little effect on the overall picture.

### 2. Choosing an SVD cut

The problematic eigenvalues are those for which

$$\frac{\lambda_n^{\text{approx}}}{\lambda_n^{\text{exact}}} < 1 - \sqrt{2/N_G} \quad (\text{D7})$$

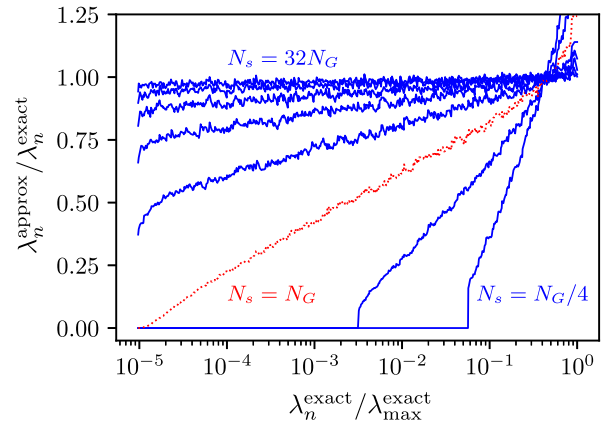


FIG. 15. The ratio of approximate to exact eigenvalues of the correlation matrix for  $N_G = 512$  correlated data points is plotted versus the size of the exact eigenvalues divided by the maximum eigenvalue. The approximate eigenvalues are determined [Eq. (D2)] from random samples of different sizes ranging (by powers of 2) from  $N_s = N_G/4$  to  $N_s = 32N_G$ . The red (dashed) line corresponds to  $N_s = N_G$ . Both sets of eigenvalues (approximate and exact) are ordered from smallest to largest.

since, on average, individual terms in  $\chi^2(\mathbf{p})$  should contribute approximately  $1 \pm \sqrt{2/N_G}$  to the total (based on the width of the  $\chi^2$  distribution). Following [91,92], we deal with these eigenvalues by introducing a cutoff  $\kappa$  such that eigenvalues smaller than  $\kappa\lambda_{\max}$  are replaced with  $\kappa\lambda_{\max}$ , where  $\lambda_{\max}$  is the largest eigenvalue:

$$\lambda_n \rightarrow \max(\lambda_n, \kappa\lambda_{\max}). \quad (\text{D8})$$

Tuning  $\kappa$  appropriately, this replacement increases the underestimated eigenvalues to a value that is at least as large as the exact eigenvalue (and probably a lot larger). Unlike in [91,92], we do not renormalize the eigenvalues to preserve the trace of the modified matrix (see below).

We need curves such as those in Fig. 15 to set  $\kappa$ , but we do not know the exact eigenvalues in real applications. An approximate curve can be generated by comparing the eigenvalues of bootstrapped copies of our simulation results  $\{\mathbf{G}^{(s)}\}$  with the eigenvalues from the full ensemble. Each bootstrapped copy has  $N_s$  samples, as the original ensemble. In this analysis, bootstrapped eigenvalues play the role of the approximate eigenvalues above, while eigenvalues computed directly from ensemble  $\{\mathbf{G}^{(s)}\}$  now play the role of the exact eigenvalues (since they specify the underlying distribution for the bootstrapped copies).

Ratios of these eigenvalues are plotted in Fig. 16 (blue points) for examples with  $N_G = 50$  (top panel) and  $N_G = 500$  (bottom panel) data points, and  $N_s = 4N_G$  random samples for each data point. The error bars show the spread of values across the different bootstrapped copies. These points give us an approximate curve for  $\lambda_n/\lambda_n^{\text{exact}}$ , from which we can determine an SVD cutoff.

The ensembles used in these examples were generated from a known distribution, so in this case we know the correct curve for  $\lambda_n^{\text{approx}}/\lambda_n^{\text{exact}}$ —that is, the ratio of the eigenvalues from the original ensemble to the exact eigenvalues from the underlying distribution. The (solid) red line in the plots shows this curve; it agrees well with the bootstrap estimates.

The vertical (dotted) red lines in each figure show the position where the ratio curves intersect with  $1 - \sqrt{2/N_G}$  [bottom dotted line, see Eq. (D7)]. We set the SVD cutoff at this point in each case. Fitting these data we find that

$$\chi^2/N_G = \begin{cases} 0.90 & \text{with no SVD cut} \\ 0.82 & \text{with } \kappa = 0.00003 \end{cases} \quad (\text{D9})$$

for  $N_G = 50$ , showing that the SVD cut has a minimal effect (as expected), while for  $N_G = 500$  we have

$$\chi^2/N_G = \begin{cases} 1.30 & \text{with no SVD cut} \\ 0.41 & \text{with } \kappa = 0.025, \end{cases} \quad (\text{D10})$$

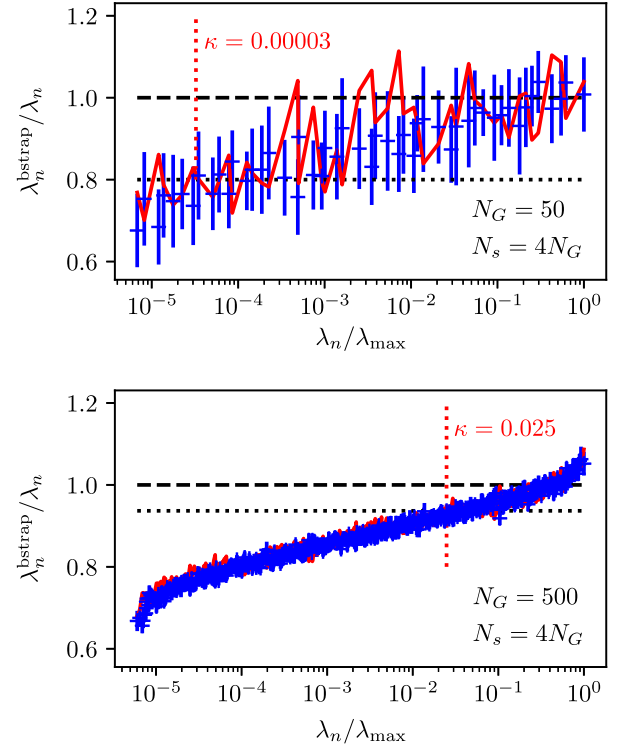


FIG. 16. Correlation-matrix eigenvalues computed from  $N_s$  random samples of  $N_G$  correlated data points are compared with eigenvalues computed from bootstrapped copies of the random sample. Results are shown for  $N_G = 50$  (top) and  $N_G = 500$  (bottom), with  $N_s = 4N_G$  in each case. The blue data points are ratios of bootstrapped eigenvalues to eigenvalues from the random sample itself; the error bars show the spread across different bootstrapped copies. The solid red line (mostly hidden in the bottom panel) shows ratios of eigenvalues from the random sample to those from the underlying distribution used to generate the random sample. The locations of the SVD cuts  $\kappa$  are shown by the vertical dashed red lines.

which shows that the SVD cut is essential since  $\chi^2/N_G = 1.30$  is much too large for  $N_G = 500$ —it corresponds to a  $p$  value of order  $3 \times 10^{-5}$ .

### 3. Conceptual framework

The nature of the SVD modification can be understood by representing the ensemble-averaged data as a vector of Gaussian random variables,

$$\mathbf{G} = \tilde{\mathbf{G}} + \delta\mathbf{G}, \quad (\text{D11})$$

where

$$\delta\mathbf{G} \equiv \sum_{n=1}^{N_G} z_n \sqrt{\lambda_n} \mathbf{D}\mathbf{v}_n, \quad (\text{D12})$$

and the uncorrelated random variables  $z_n$  satisfy

$$\langle z_n \rangle = 0, \quad \langle z_n z_m \rangle = \delta_{nm}. \quad (\text{D13})$$

Here  $\delta\mathbf{G}$  represents the uncertainty associated with the ensemble average,

$$\langle \delta\mathbf{G}\delta\mathbf{G}^T \rangle = \sum_n \lambda_n \mathbf{D}\mathbf{v}_n \mathbf{v}_n^T \mathbf{D} = M_{\text{cov}}. \quad (\text{D14})$$

The effect of the SVD cut is to add more uncertainty,  $\delta\mathbf{G}^{\text{SVD}}$ ,

$$\mathbf{G} \rightarrow \bar{\mathbf{G}} + \delta\mathbf{G} + \delta\mathbf{G}_{\text{SVD}}, \quad (\text{D15})$$

where

$$\delta\mathbf{G}_{\text{SVD}} \equiv \sum_{\lambda_n < \kappa\lambda_{\text{max}}} \tilde{z}_n \sqrt{\kappa\lambda_{\text{max}} - \lambda_n} \mathbf{D}\mathbf{v}_n, \quad (\text{D16})$$

and  $\tilde{z}_n$  are new random variables with zero mean and unit covariance matrix. Then

$$\begin{aligned} & \langle (\delta\mathbf{G} + \delta\mathbf{G}_{\text{SVD}})(\delta\mathbf{G} + \delta\mathbf{G}_{\text{SVD}})^T \rangle \\ &= \langle \delta\mathbf{G}\delta\mathbf{G}^T \rangle + \langle \delta\mathbf{G}_{\text{SVD}}\delta\mathbf{G}_{\text{SVD}}^T \rangle \\ &= \sum_n \max(\lambda_n, \kappa\lambda_{\text{max}}) \mathbf{D}\mathbf{v}_n \mathbf{v}_n^T \mathbf{D} \end{aligned} \quad (\text{D17})$$

is the SVD-modified covariance data. The SVD noise discussed above is a random sample drawn from the distribution described by  $\delta\mathbf{G}_{\text{SVD}}$ .

These formulas underscore the fact that introducing an SVD cut is a conservative move: it always increases the uncertainties in the data. This would not necessarily be the case if we renormalized the eigenvalues after introducing the SVD cut, as is done in [91,92]. In practice, however, the difference between the two approaches is small.

Finally note that another option in an SVD analyses is to discard modes below the cutoff. This corresponds to setting  $\lambda_n = \infty$  for these modes, which is much larger than is reasonable, much too conservative. We find that fits are more accurate and more stable using the prescription outlined above.

#### 4. Goodness of fit

Note that  $\chi^2/N_G = 0.41$  in Eq. (D10) is much smaller than expected for  $N_G = 500$ : one expects 1.00(6) instead. The small value arises because random fluctuations in  $\bar{\mathbf{G}}$  are characteristic of the uncertainties in  $\delta\mathbf{G}$ , but not those in  $\delta\mathbf{G}_{\text{SVD}}$ . We can demonstrate this by adding a random sample to  $\bar{\mathbf{G}}$  drawn from the distribution specified by  $\delta\mathbf{G}_{\text{SVD}}$ ,

$$\delta\mathbf{G}_{\text{SVD}} \rightarrow \text{sample}(\delta\mathbf{G}_{\text{SVD}}) + \delta\mathbf{G}_{\text{SVD}}, \quad (\text{D18})$$

and refitting. In the case of Eq. (D10), a typical fit with SVD noise gives  $\chi^2/N_G$  increases to 0.96, which is consistent with expectations.

Parenthetically, we note that overly broad priors—for example,  $0 \pm 10$  for a set of parameters that are all order 1—can also result in a small  $\chi^2$ . This situation is addressed in a similar fashion, by replacing the prior distribution  $\mathbf{P}$ :

$$\mathbf{P} \equiv \bar{\mathbf{P}} + \delta\mathbf{P} \rightarrow \bar{\mathbf{P}} + \text{sample}(\delta\mathbf{P}) + \delta\mathbf{P}. \quad (\text{D19})$$

A good fit should have  $\chi^2/N_G \approx 1 \pm \sqrt{2/N_G}$  when both SVD and prior noise is included, and the fit results should agree (within errors) with the results without noise.

A more direct test of a fitting protocol than adding extra noise is to replace the fit data [Eq. (D11)] with simulated data,

$$\mathbf{G}_{\text{sim}} \equiv \mathbf{G}(\mathbf{p}_{\text{sim}}) + \text{sample}(\delta\mathbf{G}) + \delta\mathbf{G}, \quad (\text{D20})$$

which has the same covariance matrix (from  $\delta\mathbf{G}$ ) as the real data, but whose mean is a random sample drawn from a distribution whose mean is known  $[\mathbf{G}(\mathbf{p}_{\text{sim}})]$ . A good fit to this simulated data should give best-fit results for the parameters that agree with  $\mathbf{p}_{\text{sim}}$  to within errors. An obvious choice for the simulation parameters  $\mathbf{p}_{\text{sim}}$  are the best-fit results obtained when fitting the real data.

- 
- [1] C. T. H. Davies *et al.* (HPQCD, UKQCD, MILC, and Fermilab Collaborations), High Precision Lattice QCD Confronts Experiment, *Phys. Rev. Lett.* **92**, 022001 (2004).  
 [2] R. J. Dowdall, C. T. H. Davies, R. R. Horgan, C. J. Monahan, and J. Shigemitsu (HPQCD Collaboration), B-Meson Decay Constants from Improved Lattice Non-relativistic QCD with Physical u, d, s, and c Quarks, *Phys. Rev. Lett.* **110**, 222003 (2013).

- [3] E. Dalgic, A. Gray, E. Gamiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu, H. Trotter, and M. Wingate (HPQCD Collaboration),  $B_s^0 - \bar{B}_s^0$  mixing parameters from unquenched lattice QCD, *Phys. Rev. D* **76**, 011501 (2007).  
 [4] E. Gamiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu, and M. Wingate (HPQCD Collaboration), Neutral  $B$  meson mixing in unquenched lattice QCD, *Phys. Rev. D* **80**, 014503 (2009).



- [5] A. Bazavov, C. Bernard, C. M. Bouchard, C. DeTar, M. Di Pierro *et al.*, Neutral B-meson mixing from three-flavor lattice QCD: Determination of the SU(3)-breaking ratio  $\xi$ , *Phys. Rev. D* **86**, 034503 (2012).
- [6] Y. Aoki, T. Ishikawa, T. Izubuchi, C. Lehner, and A. Soni, Neutral B meson mixings and B meson decay constants with static heavy and domain-wall light quarks, *Phys. Rev. D* **91**, 114505 (2015).
- [7] N. Carrasco *et al.* (ETM Collaboration), B-physics from  $N_f = 2$  tmQCD: The Standard Model and beyond, *J. High Energy Phys.* **03** (2014) 016.
- [8] A. Bazavov *et al.* (Fermilab Lattice and MILC),  $B_{(s)}^0$ -mixing matrix elements from lattice QCD for the Standard Model and beyond, *Phys. Rev. D* **93**, 113016 (2016).
- [9] P. A. Boyle, L. Del Debbio, N. Garron, A. Juttner, A. Soni, J. T. Tsang, and O. Witzel (RBC/UKQCD Collaboration), SU(3)-breaking ratios for  $D_{(s)}$  and  $B_{(s)}$  mesons, [arXiv:1812.08791](https://arxiv.org/abs/1812.08791).
- [10] F. Gabbiani, E. Gabrielli, A. Masiero, and L. Silvestrini, A Complete analysis of FCNC and CP constraints in general SUSY extensions of the Standard Model, *Nucl. Phys.* **B477**, 321 (1996).
- [11] M. Tanabashi *et al.* (Particle Data Group), Review of particle physics, *Phys. Rev. D* **98**, 030001 (2018).
- [12] B. Chakraborty, C. T. H. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile (HPQCD Collaboration), High-precision quark masses and QCD coupling from  $n_f = 4$  lattice QCD, *Phys. Rev. D* **91**, 054508 (2015).
- [13] A. Bazavov *et al.*,  $B$ - and  $D$ -meson leptonic decay constants from four-flavor lattice QCD, *Phys. Rev. D* **98**, 074512 (2018).
- [14] C. Monahan, E. Gamiz, R. Horgan, and J. Shigemitsu, Matching lattice and continuum four-fermion operators with nonrelativistic QCD and highly improved staggered quarks, *Phys. Rev. D* **90**, 054015 (2014).
- [15] B. Colquhoun, C. T. H. Davies, R. J. Dowdall, J. Kettle, J. Koponen, G. P. Lepage, and A. T. Lytle (HPQCD Collaboration), B-meson decay constants: a more complete picture from full lattice QCD, *Phys. Rev. D* **91**, 114509 (2015).
- [16] M. Beneke, G. Buchalla, C. Greub, A. Lenz, and U. Nierste, Next-to-leading order QCD corrections to the lifetime difference of  $B(s)$  mesons, *Phys. Lett. B* **459**, 631 (1999).
- [17] C. Monahan, J. Shigemitsu, and R. Horgan, Matching lattice and continuum axial-vector and vector currents with non-relativistic QCD and highly improved staggered quarks, *Phys. Rev. D* **87**, 034017 (2013).
- [18] A. J. Buras, S. Jager, and J. Urban, Master formulae for  $\Delta F = 2$  NLO QCD factors in the Standard Model and beyond, *Nucl. Phys.* **B605**, 600 (2001).
- [19] D. Becirevic, V. Gimenez, G. Martinelli, M. Papinutto, and J. Reyes, B parameters of the complete set of matrix elements of  $\Delta B = 2$  operators from the lattice, *J. High Energy Phys.* **04** (2002) 025.
- [20] M. Ciuchini, E. Franco, V. Lubicz, F. Mescia, and C. Tarantino, Lifetime differences and CP violation parameters of neutral B mesons at the next-to-leading order in QCD, *J. High Energy Phys.* **08** (2003) 031.
- [21] B. Chakraborty, C. T. H. Davies, G. C. Donald, J. Koponen, and G. P. Lepage (HPQCD Collaboration), Nonperturbative comparison of clover and highly improved staggered quarks in lattice QCD and the properties of the  $\phi$  meson, *Phys. Rev. D* **96**, 074502 (2017).
- [22] A. Bazavov *et al.* (MILC Collaboration), Scaling studies of QCD with the dynamical HISQ action, *Phys. Rev. D* **82**, 074501 (2010).
- [23] A. Bazavov *et al.* (MILC Collaboration), Lattice QCD ensembles with four flavors of highly improved staggered quarks, *Phys. Rev. D* **87**, 054505 (2013).
- [24] R. J. Dowdall *et al.* (HPQCD Collaboration), The Upsilon spectrum and the determination of the lattice spacing from lattice QCD including charm quarks in the sea, *Phys. Rev. D* **85**, 054509 (2012).
- [25] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trotter, and K. Wong (HPQCD Collaboration), Highly improved staggered quarks on the lattice, with applications to charm physics, *Phys. Rev. D* **75**, 054502 (2007).
- [26] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD Collaboration), High Precision Determination of the  $\pi$ , K, D and  $D_s$  Decay Constants from Lattice QCD, *Phys. Rev. Lett.* **100**, 062002 (2008).
- [27] C. T. H. Davies, C. McNeile, E. Follana, G. P. Lepage, H. Na, and J. Shigemitsu (HPQCD Collaboration), Update: Precision  $D_s$  decay constant from full lattice QCD using very fine lattices, *Phys. Rev. D* **82**, 114504 (2010).
- [28] G. C. Donald, C. T. H. Davies, R. J. Dowdall, E. Follana, K. Hornbostel, J. Koponen, G. P. Lepage, and C. McNeile (HPQCD Collaboration), Precision tests of the  $J/\psi$  from full lattice QCD: Mass, leptonic width and radiative decay rate to  $\eta_c$ , *Phys. Rev. D* **86**, 094501 (2012).
- [29] G. P. Lepage, L. Magnea, C. Nakhleh, U. Magnea, and K. Hornbostel, Improved nonrelativistic QCD for heavy quark physics, *Phys. Rev. D* **46**, 4052 (1992).
- [30] G. P. Lepage and P. B. Mackenzie, On the viability of lattice perturbation theory, *Phys. Rev. D* **48**, 2250 (1993).
- [31] J. O. Daldrop, C. T. H. Davies, and R. J. Dowdall (HPQCD Collaboration), Prediction of the Bottomonium D-Wave Spectrum from Full Lattice QCD, *Phys. Rev. Lett.* **108**, 102003 (2012).
- [32] R. J. Dowdall, C. T. H. Davies, T. Hammant, R. R. Horgan, and C. Hughes (HPQCD Collaboration), Bottomonium hyperfine splittings from lattice nonrelativistic QCD including radiative and relativistic corrections, *Phys. Rev. D* **89**, 031502 (2014); **92**, 039904 (2015).
- [33] R. J. Dowdall, C. T. H. Davies, T. C. Hammant, and R. R. Horgan (HPQCD Collaboration), Precise heavy-light meson masses and hyperfine splittings from lattice QCD including charm quarks in the sea, *Phys. Rev. D* **86**, 094510 (2012).
- [34] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD Collaboration), Precise determination of the lattice spacing in full lattice QCD, *Phys. Rev. D* **81**, 034506 (2010).
- [35] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile (HPQCD Collaboration),  $V_{us}$  from  $\pi$  and K decay constants in full lattice QCD with physical u, d, s and c quarks, *Phys. Rev. D* **88**, 074504 (2013).

- [36] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage (HPQCD Collaboration), High-precision  $c$  and  $b$  masses, and QCD coupling from current-current correlators in lattice and continuum QCD, *Phys. Rev. D* **82**, 034512 (2010).
- [37] N. Garron, R. J. Hudspith, and A. T. Lytle (RBC/UKQCD Collaboration), Neutral Kaon mixing beyond the Standard Model with  $n_f = 2 + 1$  Chiral fermions part 1: Bare matrix elements and physical results, *J. High Energy Phys.* **11** (2016) 001.
- [38] A. G. Grozin, T. Mannel, and A. A. Pivovarov, Towards a next-to-next-to-leading order analysis of matching in  $B^0 - \bar{B}^0$  mixing, *Phys. Rev. D* **96**, 074032 (2017).
- [39] D. King, A. Lenz, and T. Rauh,  $B_s$  mixing observables and  $-V_{td}/V_{ts}$  from sum rules, *J. High Energy Phys.* **05** (2019) 034.
- [40] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage (HPQCD Collaboration), High-precision  $f_{B_s}$  and HQET from relativistic lattice QCD, *Phys. Rev. D* **85**, 031503 (2012).
- [41] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage (HPQCD Collaboration), Heavy meson masses and decay constants from relativistic heavy quarks in full lattice QCD, *Phys. Rev. D* **86**, 074503 (2012).
- [42] C. Hughes, C. T. H. Davies, and C. J. Monahan (HPQCD Collaboration), New methods for B meson decay constants and form factors from lattice NRQCD, *Phys. Rev. D* **97**, 054509 (2018).
- [43] K. Abe *et al.* (Belle Collaboration), Improved measurement of CP-violation parameters  $\sin 2\phi(1)$  and  $-\lambda$ , B meson lifetimes, and  $B^0$ -anti- $B^0$  mixing parameter  $\Delta m$  (d), *Phys. Rev. D* **71**, 072003 (2005); **71**, 079903 (2005).
- [44] B. Aubert *et al.* (BABAR Collaboration), Measurement of the  $\bar{B}^0$  lifetime and the  $B^0 \bar{B}^0$  oscillation frequency using partially reconstructed  $\bar{B}^0 \rightarrow D^{*+} \ell^- \bar{\nu}_\ell$  decays, *Phys. Rev. D* **73**, 012004 (2006).
- [45] V. M. Abazov *et al.* (D0 Collaboration), Measurement of  $B_d$  mixing using opposite-side flavor tagging, *Phys. Rev. D* **74**, 112002 (2006).
- [46] R. Aaij *et al.* (LHCb Collaboration), A precise measurement of the  $B^0$  meson oscillation frequency, *Eur. Phys. J. C* **76**, 412 (2016).
- [47] A. Abulencia *et al.* (CDF Collaboration), Observation of  $B_s^0 - \bar{B}_s^0$  Oscillations, *Phys. Rev. Lett.* **97**, 242003 (2006).
- [48] R. Aaij *et al.* (LHCb Collaboration), Measurement of the  $B_s^0 - \bar{B}_s^0$  oscillation frequency  $\Delta m_s$  in  $B_s^0 \rightarrow D_s^-(3)\pi$  decays, *Phys. Lett. B* **709**, 177 (2012).
- [49] R. Aaij *et al.* (LHCb Collaboration), Precision measurement of the  $B_s^0 - \bar{B}_s^0$  oscillation frequency with the decay  $B_s^0 \rightarrow D_s^- \pi^+$ , *New J. Phys.* **15**, 053021 (2013).
- [50] R. Aaij *et al.* (LHCb Collaboration), Observation of  $B_s^0 - \bar{B}_s^0$  mixing and measurement of mixing frequencies using semileptonic B decays, *Eur. Phys. J. C* **73**, 2655 (2013).
- [51] R. Aaij *et al.* (LHCb Collaboration), Precision Measurement of CP Violation in  $B_s^0 \rightarrow J/\psi K^+ K^-$  decays, *Phys. Rev. Lett.* **114**, 041801 (2015).
- [52] T. Inami and C. S. Lim, Effects of superheavy quarks and leptons in low-energy weak processes  $k(L) \rightarrow \mu$  anti- $\mu$ ,  $K \rightarrow \pi$  neutrino anti-neutrino and  $K^0 \leftrightarrow \text{anti-}K^0$ , *Prog. Theor. Phys.* **65**, 297 (1981); **65**, 1772 (1981).
- [53] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, Weak decays beyond leading logarithms, *Rev. Mod. Phys.* **68**, 1125 (1996).
- [54] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the top quark mass using proton-proton data at  $\sqrt{s} = 7$  and 8 TeV, *Phys. Rev. D* **93**, 072004 (2016).
- [55] M. Aaboud *et al.* (ATLAS Collaboration), Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  dilepton channel from  $\sqrt{s} = 8$  TeV ATLAS data, *Phys. Lett. B* **761**, 350 (2016).
- [56] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the top quark mass with lepton + jets final states using p p collisions at  $\sqrt{s} = 13$  TeV, *Eur. Phys. J. C* **78**, 891 (2018).
- [57] P. Marquard, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, Quark Mass Relations to Four-Loop Order in Perturbative QCD, *Phys. Rev. Lett.* **114**, 142002 (2015).
- [58] A. J. Buras, M. Jamin, and P. H. Weisz, Leading and next-to-leading QCD Corrections to  $\epsilon$  parameter and  $B^0 - \bar{B}^0$  mixing in the presence of a heavy top quark, *Nucl. Phys. B* **347**, 491 (1990).
- [59] J. Charles, A. Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos (CKMfitter Group), CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories, *Eur. Phys. J. C* **41**, 1 (2005).
- [60] M. Bona *et al.* (UTfit Collaboration), The unitarity triangle fit in the Standard Model and hadronic parameters from lattice QCD: A reappraisal after the measurements of  $\Delta M_s$  and  $\text{BR}(B \rightarrow \tau \nu_\tau)$ , *J. High Energy Phys.* **10** (2006) 081.
- [61] CKMfitter, <http://ckmfitter.in2p3.fr> (2018).
- [62] UTfit, <http://www.utfit.org/UTfit> (2018).
- [63] M. Blanke and A. J. Buras, Universal unitarity triangle 2016 and the tension between  $\Delta M_{s,d}$  and  $\epsilon_K$  in CMFV models, *Eur. Phys. J. C* **76**, 197 (2016).
- [64] L. Di Luzio, M. Kirk, and A. Lenz, Updated  $B_s$ -mixing constraints on new physics models for  $b \rightarrow s \ell^+ \ell^-$  anomalies, *Phys. Rev. D* **97**, 095035 (2018).
- [65] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, and M. Steinhauser,  $B_{s,d} \rightarrow l^+ l^-$  in the Standard Model with Reduced Theoretical Uncertainty, *Phys. Rev. Lett.* **112**, 101801 (2014).
- [66] A. J. Buras, Relations between  $\Delta M(s, d)$  and  $B(s, d) \rightarrow \mu \bar{\mu}$  in models with minimal flavor violation, *Phys. Lett. B* **566**, 115 (2003).
- [67] K. De Bruyn, R. Fleischer, R. Knegjens, P. Koppenburg, M. Merk, A. Pellegrino, and N. Tuning, Probing New Physics via the  $B_s^0 \rightarrow \mu^+ \mu^-$  Effective Lifetime, *Phys. Rev. Lett.* **109**, 041801 (2012).
- [68] Y. Amhis *et al.* (HFLAV Collaboration), Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of summer 2016, *Eur. Phys. J. C* **77**, 895 (2017).
- [69] HFLAV, <http://www.slac.stanford.edu/xorg/hfag> (2018).
- [70] M. Beneke, C. Bobeth, and R. Szafron, Enhanced Electromagnetic Correction to the Rare B-Meson Decay  $B_{s,d} \rightarrow \mu^+ \mu^-$ , *Phys. Rev. Lett.* **120**, 011801 (2018).
- [71] V. Khachatryan *et al.* (CMS and LHCb Collaborations), Observation of the rare  $B_s^0 \rightarrow \mu^+ \mu^-$  decay from the combined analysis of CMS and LHCb data, *Nature (London)* **522**, 68 (2015).
- [72] R. Aaij *et al.* (LHCb Collaboration), Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  Branching Fraction and Effective Lifetime and

- Search for  $B^0 \rightarrow \mu^+ \mu^-$  Decays, *Phys. Rev. Lett.* **118**, 191801 (2017).
- [73] M. Aaboud *et al.* (ATLAS Collaboration), Study of the rare decays of  $B_s^0$  and  $B^0$  mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector, *J. High Energy Phys.* **04** (2019) 098.
- [74] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the  $B_s^0 \rightarrow \mu^+ \mu^-$  Branching Fraction and Search for  $B^0 \rightarrow \mu^+ \mu^-$  with the CMS Experiment, *Phys. Rev. Lett.* **111**, 101804 (2013).
- [75] A. Lenz and U. Nierste, Theoretical update of  $B_s - \bar{B}_s$  mixing, *J. High Energy Phys.* **06** (2007) 072.
- [76] C. T. H. Davies, J. Harrison, G. P. Lepage, C. J. Monahan, J. Shigemitsu, and M. Wingate, Lattice QCD matrix elements for the  $B_s^0 - \bar{B}_s^0$  width difference beyond leading order, [arXiv:1910.00970](https://arxiv.org/abs/1910.00970).
- [77] A. T. Lytle, C. T. H. Davies, D. Hatton, G. P. Lepage, and C. Sturm (HPQCD Collaboration), Determination of quark masses from  $\mathbf{n}_f = 4$  lattice QCD and the RI-SMOM intermediate scheme, *Phys. Rev. D* **98**, 014513 (2018).
- [78] G. P. Lepage, B. Clark, C. T. H. Davies, K. Hornbostel, P. B. Mackenzie, C. Morningstar, and H. Trotter, Constrained curve fitting, *Nucl. Phys. B, Proc. Suppl.* **106**, 12 (2002).
- [79] C. Bouchard (private communication).
- [80] K. Hornbostel, G. P. Lepage, C. T. H. Davies, R. J. Dowdall, H. Na, and J. Shigemitsu (HPQCD Collaboration), Fast fits for lattice QCD Correlators, *Phys. Rev. D* **85**, 031504 (2012).
- [81] C. Bernard (MILC Collaboration), Neutral B mixing in staggered Chiral perturbation theory, *Phys. Rev. D* **87**, 114503 (2013).
- [82] W. Detmold and C. J. D. Lin, Matrix elements of the complete set of  $\Delta B = 2$  and  $\Delta C = 2$  operators in heavy meson chiral perturbation theory, *Phys. Rev. D* **76**, 014501 (2007).
- [83] B. Colquhoun, R. J. Dowdall, J. Koponen, C. T. H. Davies, and G. P. Lepage (HPQCD),  $B \rightarrow \pi \ell \nu$  at zero recoil from lattice QCD with physical u/d quarks, *Phys. Rev. D* **93**, 034502 (2016).
- [84] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), B- and D-meson decay constants from three-flavor lattice QCD, *Phys. Rev. D* **85**, 114506 (2012).
- [85] W. Detmold, C. J. D. Lin, and S. Meinel, Axial Couplings and Strong Decay Widths of Heavy Hadrons, *Phys. Rev. Lett.* **108**, 172003 (2012).
- [86] F. Bernardoni, J. Bulava, M. Donnellan, and R. Sommer (ALPHA Collaboration), Precision lattice QCD computation of the  $B^* B \pi$  coupling, *Phys. Lett. B* **740**, 278 (2015).
- [87] J. M. Flynn, P. Fritzsch, T. Kawanai, C. Lehner, B. Samways, C. T. Sachrajda, R. S. Van de Water, and O. Witzel (RBC, UKQCD), The  $B^* B \pi$  coupling using relativistic heavy quarks, *Phys. Rev. D* **93**, 014510 (2016).
- [88] B. Grinstein, E. E. Jenkins, A. V. Manohar, M. J. Savage, and M. B. Wise, Chiral perturbation theory for  $f_{D(s)}/f_D$  and  $B_{B(s)}/B_B$ , *Nucl. Phys. B* **380**, 369 (1992).
- [89] D. Becirevic, S. Fajfer, and J. F. Kamenik, Chiral behavior of the  $B^0(d, s) - \bar{B}^0(d, s)$  mixing amplitude in the standard model and beyond, *J. High Energy Phys.* **06** (2007) 003.
- [90] D. Arndt and C. J. D. Lin, Heavy meson chiral perturbation theory in finite volume, *Phys. Rev. D* **70**, 014503 (2004).
- [91] C. Michael, Fitting correlated data, *Phys. Rev. D* **49**, 2616 (1994).
- [92] C. Michael and A. McKerrell, Fitting correlated hadron mass spectrum data, *Phys. Rev. D* **51**, 3745 (1995).